QUANTITATIVE GENETIC MAPPING OF LIFE HISTORY TRAITS IN *Drosophila melanogaster*

By

Casey Lee M^{c}Neil

Submitted to the graduate degree program in Molecular Biosciences
and the Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy.

_____
Chairperson – Stuart Macdonald

_____
Brian Ackley

_____
Matthew Buechner

_____
Erik Lundquist

_____
Robert Ward

_____
John Kelly

Date Defended: July 23, 2012

The Dissertation Committee for Casey Lee McNeil
certifies that this is the approved version of the following dissertation:

QUANTITATIVE GENETIC MAPPING OF LIFE HISTORY TRAITS IN *Drosophila melanogaster*

_____
Chairperson – Stuart Macdonald

Date approved: July 23, 2012

Abstract

While it has long been established that populations of animals harbor substantial natural genetic variation for life history traits, an understanding of the location, effect, and frequency of naturally occurring alleles has been elusive. This study uses the elite model system *Drosophila melanogaster* to perform quantitative genetic mapping on two important life history traits: the morphology of the posterior lobe of the genital arch and the length of time flies resist death due to starvation stress. Experiments on the posterior lobe identify multiple quantitative trait loci (QTL) that influence its shape and reveal a small number of strong candidate genes for future study. Studying starvation stress we identify many small effect QTL that generally act in a sex-specific manner. Using a series of crosses we both identify many new cross-specific QTL while replicating all QTL originally identified among inbred lines in outbred genetic backgrounds. These results indicate that inbreeding depression is likely not playing a major role in genetic mapping results obtained with inbred animals. These studies identify an exceedingly complex genetic architecture for starvation stress resistance in *Drosophila* that may include additive, dominant, and epistatic alleles acting to influence this important life history trait.

Acknowledgements

To my parents, Daniel L. and Donna M. McNeil, who gave Dan and I a tremendous start in life, and who always encouraged me to pursue my dreams.

To my twin brother, Daniel W. McNeil, who modeled to me how to work selflessly and how to be an outstanding brother.

To Steven R. Hoofer, who taught me how to handle being both a scientist and a man of faith.

To Clint L. Bain, whose work ethic and ability are second to none, and who made me a better student, teacher, and person. Take 'er easy, Dude.

To Stuart J. Macdonald, who taught me everything I know about doing big experiments, analyzing data, and being a professional scientist.

To John C. Mullican, whose dedication to all of his students inspired me to become a Professor. You challenged me when I didn't know how to learn and fostered my ability once I discovered it. I am eternally grateful for your support and friendship.

Finally, to my wife, Ashlee, you are my best friend and the love of my life. I enjoy every day we have together, and look forward to so many more!

Table of Contents

List of Figures

Chapter 4

List of Tables

Introduction

Natural genetic variation influencing organismal fitness is largely uncharacterized, especially at a molecular level, and interest in these phenotypes has been outweighed by technical difficulties surrounding their study (e.g. limited heritability and error in phenotyping assays). In addition, the discovery of specific alleles that generate quantitative genetic variation among animals in nature has generally been hampered by their complex genetic architecture. Linkage mapping studies attempting to map QTL (quantitative trait loci) have traditionally suffered from the inability to precisely resolve QTL (Mackay, 2001), making the elucidation of potentially causative genes and alleles both time consuming and expensive. Moreover, since traditional QTL mapping approaches utilized only two parental lines, essential population genetic parameters --notably allelic frequency information -- of any mapped QTL remained unknown.

Recent development of multi-parental, high-resolution mapping panels, such as the mouse collaborative cross (Aylor, *et al*. 2011), *Arabidopsis* MAGIC lines (Kover *et al*. 2009), the maize NAM panel (Buckler, *et al*. 2009), and the *Drosophila* Synthetic Population Resource (DSPR) (King, *et al*. 2012), allow mapping QTL of even small effect sizes to narrow genomic intervals, while estimating the frequency of causative alleles among natural populations of animals. These approaches, taking advantage of high-throughput sequencing of parental lines and modern genotyping technologies for the mapping panel, simply requires a commitment to phenotype vast numbers of experimental individuals. We did exactly that by phenotyping 1725 recombinant inbred lines (RIL) of the

DSPR for the time of death by starvation. We utilize this mapping data to attempt to identify naturally occurring alleles that cause quantitative differences in starvation stress resistance in *Drosophila*.

Utilizing our ability to map QTL of small effect, we endeavoured to characterize the genetic architecture of starvation resistance across several genetic backgrounds. Phenotyping the F_1 progeny of three independent crosses, we recapitulate many of the QTL identified by direct screening in homozygous RILs and discover many novel, cross specific QTL. Our results not only suggest an exceedingly complex genetic architecture for starvation resistance, but also point toward the importance of characterizing quantitative traits in multiple genetic backgrounds.

Life history traits such as stress resistance are not the only phenotypes to exert a direct impact on organismal fitness in nature. Many morphological structures also play a role in the ability of an individual to survive and reproduce. One rapidly developing morphological structure that may directly affect fitness is the posterior lobe of the genital arch of *Drosophila.* The posterior lobe is a novel cuticular structure limited to males of the *D. melanogaster* species complex -- *D. melanogaster, D. mauritiana*, *D. sechellia,* and *D. simulans*. The posterior lobe appears to play an essential role in copulation in these species: During physical coupling, the posterior lobe is inserted under the ninth tergite of the female, helping to establish a robust genital coupling between the male and female *Drosophila* (Jagadeeshan and Singh, 2006).

While several genetic mapping experiments have succeeded in identifying genomic intervals that contribute to phenotypic differences in interspecific hybrid crosses (Liu, *et al*. 1996, Laurie, *et al*. 1997, Macdonald and Goldstein, 1999) we are the first to investigate morphological variation within *D. melanogaster.* We utilize both coarse (F_2) and fine (F_{17}) mapping panels to identify at least three QTL contributing to intraspecific variation in the shape of the posterior lobe (McNeil, *et al*. 2011). Comparing our QTL mapping data with genes found to be differentially expressed between male and female *Drosophila* genital imaginal discs (the primordia of the adult genitalia) at three developmental time points (Chatterjee, *et al*. 2011), we find a significant association between the two datasets, succeeding in defining a small number of plausible candidate genes underlying posterior lobe morphology. This list of putatively causative genes gives us a number of direct targets for functional analysis and importantly, serves as a benchmark for future mapping studies. Identifying QTL or associations at or near the current mapped QTL will bolster these results and increase the confidence of investigators seeking to identify the precise alleles influencing the morphology of the male genitalia in *Drosophila.*

Literature Cited

Aylor, D.L., Valdar, W., Foulds-Mathes, W., Buus, R.J., Verdugo, R.A., Baric, R.S., Ferris, M.T., Frelinger, J.A., Heise, M., Frieman, M.B., *et al*. 2011. Genetic analysis of complex traits in the emerging Collaborative Cross. *Genome Res* **21:** 1213–1222.

Buckler, E.S., Holland, J.B., Bradbury, P.J., Acharya, C.B., Brown, P.J., Browne, C., Ersoz, E., Flint-Garcia, S., Garcia, A., Glaubitz, J.C., *et al*. 2009. The genetic architecture of maize flowering time. *Science* **325:** 714–718.

Chatterjee, S.S., L.D. Uppendahl, M.A. Chowdhury, P-L. Ip, and Siegal, M.L. 2011. The female-specific Doublesex isoform regulates pleiotropic transcription factors to pattern genital development in *Drosophila*. *Development* **138:** 1099–1109.

Jagadeeshan, S. and Singh, R.S. 2006. A time-sequence functional analysis of mating behaviour and genital coupling in *Drosophila*: role of cryptic female choice and male sex-drive in the evolution of genitalia. *J. Evol. Biol*. **19:** 1058–1070.

King, E.G., Merkes, C.M., McNeil, C.L., Hoofer, S.R., Sen, S., Broman, K.W., Long, A.D., and Macdonald, S.J. 2012. Genetic dissection of a model complex trait using the *Drosophila* Synthetic Population Resource. *Genome Research*

(Online before print, April 2012)

Kover, P.X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I.M., Purugganan, M.D., Durrant, C., and Mott, R. 2009. A Multiparent Advanced Generation Inter-Cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet* **5:** e1000551.

Laurie, C. C., J. R. True, J. Liu, and J. M. Mercer, 1997 An introgression analysis of quantitative trait loci that contribute to a morphological difference between *Drosophila simulans* and *D. mauritiana*. *Genetics* **145:** 339–348.

Liu, J., J. M. Mercer, L. F. Stam, G. C. Gibson, Z-B. Zeng *et al.*, 1996 Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. *Genetics* **145:** 1129–1145.

Macdonald, S. J., and D. B. Goldstein, 1999 A quantitative genetic analysis of male sexual traits distinguishing the sibling species *Drosophila simulans* and *D. sechellia*. *Genetics* **153:** 1683–1699.

Mackay, T.F. 2001. The genetic architecture of quantitative traits*. Annu Rev Genet* **35:** 303–339.

McNeil, C.L., Bain, C.L., Macdonald, S.J. 2011. Multiple Quantitative Trait Loci Influence the Shape of a Male-Specific Genital Structure in *Drosophila melanogaster*. *G3* 1(5):343-351.

Chapter II

Multiple quantitative trait loci influence the shape of a male-specific genital structure in *Drosophila melanogaster*

Published in *G3*, October 2011.

M^{c}Neil, C.L., Bain, C.L., Macdonald, S.J. 2011. Multiple Quantitative Trait Loci Influence the Shape of a Male-Specific Genital Structure in *Drosophila melanogaster*. G3 1:343-51

ABSTRACT

The observation that male genitalia diverge more rapidly than other morphological traits during evolution is taxonomically widespread, and likely due to some form of sexual selection. One way to elucidate the evolutionary forces acting on these traits is to detail the genetic architecture of variation both within and between species, a program of research that is considerably more tractable in a model system. *Drosophila melanogaster* and its sibling species *D. simulans*, *D. mauritiana*, and *D. sechellia* are morphologically distinguishable only by the shape of the posterior lobe, a male-specific elaboration of the genital arch. We extend earlier studies identifying QTL (quantitative trait loci) responsible for lobe divergence across species, and report the first genetic dissection of lobe shape variation within a species. Using an advanced intercross mapping design we identify three autosomal QTL contributing to the difference in lobe shape between a pair of *D. melanogaster* inbred lines. The QTL each contribute 4.6-10.7% to shape variation, and two show a significant epistatic interaction. Interestingly, these intraspecific QTL map to the same locations as interspecific lobe QTL, implying some shared genetic control of the trait within and between species. As a first step towards a mechanistic understanding of natural lobe shape variation, we find an association between our QTL data and a set of genes that show sex-biased expression in the developing genital imaginal disc (the precursor of the adult genitalia). These genes are good candidates to harbor naturally-segregating polymorphisms contributing to posterior lobe shape.

INTRODUCTION

There is a great deal of interest in characterizing the morphological and behavioral changes that distinguish closely-related species to understand the evolutionary processes involved in the early stages of speciation. In this context, the male genitalia of insects have come under particular scrutiny due to the observation that genital morphology is often species-specific, and can show striking diversity across related taxa that are otherwise similar in form (Eberhard 1985). Various lines of evidence point to sexual selection as a likely driver of this rapid divergence in genital morphology (Eberhard 1985, 2010; Hosken and Stockley 2004). Although the precise mechanism of sexual selection is debated, a popular hypothesis is cryptic female choice (Eberhard *et al.* 1998; Eberhard 2010). Indeed, several studies have demonstrated that variation in male fertilization success is linked to morphological variation in male genital structures (*e.g.*, Arnqvist and Danielsson 1999; Danielsson and Askenmo 1999; reviewed by Eberhard 2011). Despite interest in the variation and evolution of male genitalia, few studies have attempted to genetically dissect these traits (although see Sasabe *et al.* 2010 and Schafer *et al*. 2011). Describing the genetic loci responsible for phenotypic variation in terms of their allelic effects, population frequencies, and interactions can provide valuable information about the evolutionary forces acting on a trait (Templeton 1981; Mitchell-Olds *et al.* 2007).

Drosophila melanogaster is one of a handful of elite model genetic systems, and has been widely employed to characterize the genetic architecture of trait variation (Flint and Mackay 2009; Mackay 2010). Numerous related Drosophilid species can also be

reared easily in the laboratory, and recent large-scale sequencing efforts have generated genome sequences for many of them (Clark *et al.* 2007), increasing their utility as experimental organisms for comparative work. In addition, the three species most closely-related to *D. melanogaster* – *D. mauritiana*, *D. sechellia* and *D. simulans* – are reproductively interfertile, allowing recombinant individuals to be produced for genetic analysis of traits distinguishing the species. Thus, this species group provides an excellent platform with which to understand the forces that shape phenotypic variation.

Interestingly, all four species of the *D. melanogaster* complex are morphologically very similar except for the shape and size of an elaborate cuticular projection (the posterior lobe) on the male genital arch, a structure that is the only reliable morphological indicator of species identity (Ashburner *et al.* 2005). The posterior lobe inserts under the ninth abdominal tergite of the female during copulation (Robertson 1988), and is used by the male during mounting and the early stages of mating to maintain strong genital coupling (Jagadeeshan and Singh 2006). Although no formal association has been made between posterior lobe morphology and male mating success, the striking variation across species suggests directional, sexual selection acting on the structure. Thus, we have an opportunity to genetically dissect a rapidly-evolving, male-specific genital trait using the armamentarium of genetic tools available for the *Drosophila* experimental system.

A number of studies have examined divergence between *Drosophila* species in posterior lobe morphology using QTL (quantitative trait locus) mapping techniques. Early work confirmed that interspecific variation for the trait is polygenic, with separate

crosses between *D. simulans* and *D. melanogaster*, *D. mauritiana*, and *D. sechellia* all identifying at least one genetic factor contributing to phenotypic variation on each of the three major chromosomes (Coyne 1983; Coyne and Kreitman 1986). Subsequent work using larger panels of recombinants and genomewide sets of markers identified multiple QTL on each chromosome for the *D. simulans* × *D. mauritiana* cross (Liu *et al.* 1996; Laurie *et al.* 1997; Zeng *et al.* 2000) and the *D. simulans* × *D. sechellia* cross (Macdonald and Goldstein 1999). Collectively, these studies suggest that QTL contributing to posterior lobe divergence between species are numerous, show limited epistasis, and are predominantly additive. Strikingly, additive effects were nearly always in the same direction – substituting a *D. simulans* allele for a *D. mauritiana* or a *D. sechellia* allele at a lobe QTL always gave a more *D. simulans*-like lobe phenotype. This suggests a consistent history of strong directional selection acting on the trait during species divergence (Orr 1998).

Despite the work on interspecific variation in genital morphology between members of the *D. melanogaster* complex, no study has yet described natural genetic variation for the posterior lobe within any one of these species. If we can characterize the loci that maintain the fairly subtle lobe shape variation within a species, as well as those that influence extreme posterior lobe diversification among species, we can elucidate the relationship between intra- and interspecific genetic variation, and develop a detailed understanding of the selective forces operating on the trait. In this study we survey a series of *D. melanogaster* inbred lines and find considerable variation in posterior lobe morphology. We then carry out QTL mapping, employing an advanced

generation recombinant population to genetically dissect variation between a pair of lines that differ in posterior lobe shape.

MATERIALS AND METHODS

D. melanogaster stocks

Fifteen highly-inbred, *P*-element and *Wolbachia*-free lines were used in this study. Fourteen were obtained from stock centers, and were subjected to multiple generations of brother-sister mating prior to this study (see Table 1 of Macdonald and Long 2007). The remaining isogenic line, *Samarkand* ry^{506} (hereafter, *Sam*), which harbors a mutant eye-color allele at the third chromosome *rosy* locus, was provided by TFC Mackay and is described in Lyman *et al.* (1996).

Unless otherwise stated all flies were reared at 23°C under constant light, using 10 ml of cornmeal-molasses-yeast medium in polystyrene vials (25 × 95 mm).

Experimental flies

Survey of intraspecific variation in genital morphology: For each of the 15 strains we generated three or four replicate vials, collected males under CO_2 anesthesia, and stored them at –20°C in 1.5 ml microcentrifuge tubes until dissection. An average of 22.8 males were successfully phenotyped per strain (range = 12-42), with a mean of 6.84 per replicate vial.

F_2 coarse-mapping population: We chose a pair of strains with divergent lobe morphology – b3852 and *Sam* – and initiated multiple replicate cross vials with 10 virgin

b3852 females and 10 *Sam* males. Parental flies were removed within 48 hours to maintain a relatively constant low larval density. F_1 hybrid progeny were collected and aged in single-sex groups to ensure females were virgin, then multiple replicate intercross vials holding 10 virgin F_1 females and 10 F_1 males were set up. Again, flies were removed within 48 hours. Upon maturation, F_2 males from each replicate vial were collected and frozen as described above.

F_{17} fine-mapping population: Reciprocal crosses between b3852 and *Sam* were carried out in small polypropylene bottles (8 oz, 60 × 130 mm). Approximately 200 F_1 individuals from each reciprocal cross were mixed, and the combined population split into two fresh bottles. In the next generation, F_2 flies were combined into a single large glass bottle (64 oz), and this recombinant population was maintained at high census size with 12-13 day generations until the F_{16} generation eclosed. A large number of replicate vials were each initiated with ~20 F_{16} individuals, and flies allowed to lay eggs for 24 hours. F_{17} males were collected and frozen as described.

Phenotype acquisition

The terminalia was dissected from each experimental male, individually placed in a 0.2 ml PCR tube containing a drop of 1M KOH, and boiled for 2-5 min to dissolve unwanted connective tissue. For recombinant F_2 and F_{17} flies the remainder of the dissected animal was re-frozen for subsequent DNA extraction. The genital arch, including the paired posterior lobes and lateral plates, was then dissected out in 1M KOH and mounted on a microscope slide under a coverslip in a small drop of Aqua-Mount (Lerner #13800 via VWR #41799-008). Slides were left overnight at 40-45°C on a slide warmer

with a ~4 g weight pushing the coverslip down, and the next day a TIFF image of each slide-mounted posterior lobe was captured at 400X total magnification. The dissection of an experimental individual was considered successful if at least one of the pair of posterior lobes was undamaged.

For lobe data to be comparable across genotypes lobes must be placed in a standard configuration prior to morphometric analysis, *i.e.*, all lobes should have the same handedness, orientation, and relative location. To ensure all lobes were of the same handedness, images were manually flipped such that the lateral plate (and thus the "point" of the lobe) points clockwise (refer to Figure 3). Each image was then manually outlined in *ImageJ* (http://rsb.info.nih.gov/ij/index.html) using a custom macro to automatically record a set of Cartesian coordinates defining each outline. Following previous work on the posterior lobe (Liu *et al.* 1996; Macdonald and Goldstein 1999; Masly *et al.* 2011), outlines were closed with an artificial baseline that extends from the point at which the lateral plate connects to the posterior lobe. Outlines from all lobes were subsequently oriented to make these baselines horizontal. Finally, the origin of each set of coordinates was placed at the centroid of the outline to make the locations of all lobe outlines comparable.

Due to the lack of reliable morphological landmarks on the posterior lobe we used elliptic Fourier analysis, EFA, to describe outline shape (Kuhl and Giardina 1982; Ferson *et al.* 1985). We applied EFA using a custom *R* script (http://www.r-project.org/). A detailed description of the methodology as applied to posterior lobe shape is provided in Liu *et al.* (1996) and Macdonald and Goldstein (1999). Briefly, elliptic Fourier functions use a parametric representation of the *x*- and *y*-projections of the outline,

treating each independently as a function of contour length. Following EFA, each outline is represented by a set of 4*n* Fourier coefficients that can reproduce the outline with arbitrary precision depending on the number of harmonics (*n*). Here we use 25 harmonics, which provides a near-perfect reconstruction of the original outline (see Figure 2 in Liu *et al.* 1996), and yields 100 coefficients per lobe. Because we had placed the outlines in a standard configuration prior to EFA, in our analyses we did not employ the coefficient normalizing functions described in Kuhl and Giardina (1982). However, we obtained practically identical QTL mapping results for the mPC1 shape measure whether or not we applied these functions (data not shown). In addition, since posterior lobe morphology is largely unaffected by variation in overall body size (Liu *et al.* 1996; Macdonald and Goldstein 1999; Shingleton *et al.* 2009; Masly *et al.* 2011) we did not seek to control for such variation, for instance by measuring wing area or tibia length.

The 100 Fourier coefficients for a subset of experimental individuals were treated as variables in a principal components analysis (PCA) to encapsulate shape variation in a small number of mathematical descriptors. Two separate PCA were carried out using the 'prcomp' *R* function, one for the species diversity experiment, and one for the mapping experiment. The species diversity PCA consisted of individuals from the set of 15 strains used to examine morphological variation within *D. melanogaster*. The mapping experiment PCA employed all mapping population flies and their progenitors (b3852, *Sam*, F_1, F_2 and F_{17}). We caution that the principal component (PC) shape descriptors may not be comparable across these two analyses, and to avoid confusion we prefix principal components derived from the species diversity PCA with an "s" (*e.g.*, sPC1), and those from the mapping experiment PCA with an "m" (*e.g.*, mPC1).

Finally, we estimated the size of each lobe as the area enclosed by the outline, lobe height (width) as the length of the vertical (horizontal) line through the centroid, and the height : width ratio (H:W) as the ratio of these two distances.

Genetic markers

Markers discriminating b3852 and *Sam* were identified by sequencing a series of 1-kb PCR fragments in both lines. SNPs were submitted to the Illumina GoldenGate assay design tool, and 96 high-scoring SNPs spread along the three major chromosomes were chosen for genotyping in our F_2 and F_{17} mapping panels (File S1). DNA was extracted from each phenotyped recombinant using the Puregene cell and tissue kit (Qiagen), resuspended in 20 µl of 1X TE, and 10 µl of diluted DNA was used for genotyping (Illumina BeadXpress platform, UC Davis Genome Center). The resulting raw intensity data was submitted to a custom set of *R* scripts to call genotypes (see Macdonald *et al.* 2005), and 87/96 SNPs yielded high-quality genotypes (X = 16, 2L = 22, 2R = 17, 3L = 12, 3R = 20). We also genotyped a single RFLP marker at the *eyeless* gene on chromosome 4 in the F_2 mapping panel. Briefly, we amplified a short PCR fragment containing a diagnostic SNP (eyeF, 5'-TGT GTG AGC AAA ATT CTC GG-3'; eyeR, 5'-GTT TCG GCA TGG TAG GAC AT-3'), digested with *Mbo*II, and genotyped by separating restriction fragments on a 2.5% agarose gel.

QTL mapping

For the recombinant flies, depending on the quality of the dissected material, phenotypes were scored on either one or both of the posterior lobes. When both lobes

were successfully imaged (153/711, or 21.5% of the recombinants) we randomly chose the phenotype from a single lobe for mapping. All QTL mapping analyses, and estimation of the genetic map from the marker genotypes, was carried out within *R/qtl* (Broman and Sen 2009). Input files are available in File S2 and S3. For the F_2, both interval mapping (IM, Lander and Botstein 1989) and composite interval mapping (CIM, Zeng 1994) were performed using the multiple imputation method of Sen and Churchill (2001) with 256 imputations. Statistical significance was determined from 1,000 permutations (Churchill and Doerge 1994), taking care to generate *X*- and autosome-specific thresholds (Broman *et al.* 2006). For the F_{17} we took a selective genotyping approach to fine-map QTL influencing the mPC1 measure of shape. Of the 344 phenotyped F_{17} males, the 47 with the lowest, most b3852-like mPC1 score, and the 47 with the highest, most *Sam*-like mPC1 score were genotyped. To minimize analytical bias associated with selective genotyping, all phenotyped F_{17} individuals were included in the QTL mapping analysis (IM using multiple imputation), with the genotypes from the non-tail individuals recorded as missing (Lander and Botstein 1989; Sen *et al.* 2005). In addition, a stratified permutation test was carried out, separately permuting phenotypes within the genotyped and ungenotyped subsets of F_{17} individuals (Manichaikul *et al.* 2007).

RESULTS AND DISCUSSION

Variation in posterior lobe morphology within *D. melanogaster*

The shape and size of the posterior lobe differs among the four members of the *melanogaster* complex of species (see Figure 1 in Liu *et al.* 1996). In addition to this dramatic interspecific variation, more subtle intraspecific variation has been noted for *D. simulans*, *D. mauritiana*, and *D. sechellia* (Liu *et al.* 1996; Macdonald and Goldstein 1999). We extend these surveys of variation to *D. melanogaster*, and score individuals from 15 inbred lines to generate a framework for understanding the basis of lobe shape and size variation in this model genetic system.

Since the posterior lobe lacks clear morphological landmarks, we quantified size and shape variation using morphometric analyses based on sets of Cartesian coordinates defining lobe outlines. An elliptic Fourier analysis (EFA) of each outline results in a series of Fourier coefficients, and a principal components analysis (PCA) of these values encapsulates variation across individuals in a series of orthogonal descriptors of shape. Figure 1 highlights shape variation along two of these descriptors, sPC1 and sPC2, that together explain >70% of lobe variation in our sample of lines. While PCA provides a convenient small set of mathematical descriptors of shape, their interpretation is difficult due to the shear complexity of the shape variation across lines (see Figure S1). Nonetheless, careful examination of the point clouds from Figure 1, along with the relevant columns from Table 1, demonstrate clustering of individuals from the same line, and clear differences among lines. For instance, lines t7 and b3846 are separated along the sPC1 axis, while lines b3870 and t0 are separated along the sPC2 axis. These results show that our morphometric descriptions of shape are robust, and allow discrimination of the different lobe shapes found in various genotypes of *D. melanogaster*.

The lines chosen for our survey were collected from sites in ten different countries, and hence capture a large swath of the cosmopolitan genetic variation in *D. melanogaster*. However, because we did not sample multiple genotypes from the same population we cannot assess relative levels of within- and between-population variation in the posterior lobe. It may be that the extent of posterior lobe variation we describe, perhaps due to some degree of local adaptation, is greater than would be observed within a single population. A more extensive survey of morphological variation, including multiple genotypes from multiple different populations is needed to address this question.

A primary goal of our survey was to identify a pair of lines that are morphologically distinct, and differ along a major axis of intraspecific phenotypic variation, to be used as the parents for a QTL mapping study. We selected lines b3852 and *Sam* for this purpose (red symbols in Figure 1). These lines have similar lobe areas, differ strongly in sPC1 (the major axis of shape variation in the diversity panel), but not in sPC2 or sPC3, with b3852 having taller, and more narrow posterior lobes than *Sam* (Table 1).

Phenotypic description of mapping population genotypes

Lines b3852 and *Sam* were intercrossed in separate experiments to generate F_2 and F_{17} males. Posterior lobe outlines from all relevant genotypes (b3852, *Sam*, F_1, F_2, and F_{17}) were processed via EFA, and the coefficients used as variables in a PCA. The top six principal components each explain >1% of the posterior lobe variation among this set of individuals: mPC1 (62.6%), mPC2 (17.1%), mPC3 (12.2%), mPC4 (2.1%), mPC5

(1.8%), and mPC6 (1.2%). In addition, the parental strains are significantly different for each of the first four principal components: mPC1 (*t*-test, $P < 1 \times 10^{-29}$), mPC2, ($P < 1 \times 10^{-4}$), mPC3 ($P = 0.003$), and mPC4 ($P < 1 \times 10^{-7}$). However, the phenotypic distributions of the parental strains fail to overlap only for mPC1 (Figure 2).

mPC1 alone appears to provide the clearest descriptor of posterior lobe morphological variation in the b3852 × *Sam* cross. This is highlighted in Figure S2, which sorts the mapping population individuals by their mPC1 score, and demonstrates a clear morphological transition from the b3852 lobe phenotype to the *Sam* lobe phenotype as mPC1 score increases. Because the correlation between lobe area and mPC1 is low (Table 2), we are able to consider lobe size and lobe shape (as measured by mPC1), as separate sources of morphological variation in this cross. Figure 2 shows the average mPC1 score in both parental lines, the F_1, and both the F_2 and F_{17} recombinant populations. As expected, the genetically variable samples show greater variation than the parentals and hybrids. In addition, the F_1 hybrid males have a mPC1 phenotype that is midway between the parental line means, suggesting the trait is largely additive (see also Figure 3).

Principal components can be difficult to interpret in terms of familiar shape concepts, and we sought to define what aspect of lobe shape mPC1 describes in this cross. We measured the height and width of each lobe as the vertical and horizontal distance through the outline centroid, respectively, and took the ratio of height : width (H:W). Figure 3 shows that H:W and mPC1 show a strong negative relationship in the parental lines and the F_1, and we found a strong negative correlation between the traits in both the F_2 and the F_{17} ($r = -0.86$ and -0.67, respectively; Table 2). Thus, this quite

crude H:W shape measure describes much of the same shape variation encapsulated by mPC1, and allows us to think of mPC1 as predominantly describing how squat or slender a lobe is.

Coarse QTL mapping

We first carried out standard F_2 QTL mapping to provide a coarse map of loci contributing to morphological variation between b3852 (tall, narrow lobe) and *Sam* (low, broad lobe). Using interval mapping (IM) on 367 F_2 individuals genotyped for a genomewide panel of markers, we identified an extremely strong QTL on chromosome 3 for mPC1 (LOD = 79.2 close to the centromere; top panel of Figure 4), and two smaller QTL near the tip of 2L (LOD = 4.9 and 3.7). These same QTL were also identified for the H:W shape measure, consistent with the strong correlation between this trait and mPC1 (Table 2). These QTL mapping analyses were conducted using the phenotypic score from just a single lobe per individual, but when we repeated the analysis and substituted data from the other lobe (if available) we identified the same QTL (Figure S3). This result was anticipated since there is a strong correlation between the mPC1 shape score for the paired lobes (r = 0.85).

Although IM has high power to simply identify QTL, it can provide unreliable estimates of the number and location of QTL (Zeng 1994; Cornforth and Long 2003). Therefore, we applied CIM to the mPC1 dataset to increase precision and further resolve QTL. Using a window size of 10 cM and 4 marker covariates we found a novel QTL on the X, a single QTL at the tip of 2L, and resolved the broad third chromosome QTL into three separate QTL (Figure 4). However, by manipulating the window size and

altering the number of markers fitted to the model, we found we could generate quite different LOD profiles, although all runs did include a large QTL interval spanning the chromosome 3 centromere.

IM was applied to all other phenotypes measured in the F_2 – lobe area, height, width, and mPC2-mPC6 – regardless of the proportion of morphological variation explained, or whether the trait discriminated the parental lines. The likelihood profiles shown in Figure S3 reveal a number of additional QTL underlying various aspects of lobe morphology. Notably, the lobe height and width LOD profiles are similar to those for mPC1 and H:W, reflecting the strong correlation between these traits (Table 2). The profiles for lobe area and mPC3 also follow each other closely, with QTL at the tip of the X and the middle of 3R, again due to a high positive correlation between the traits (Table 2).

Finally, we note that chromosome 4 failed to show a significant association with any trait tested in the F_2 panel (data not shown).

Fine-mapping mPC1 QTL

Ultimately, rather than applying additional statistical analyses to a standard F_2 dataset, the best way to improve QTL mapping resolution, generate accurate estimates of QTL effects, and promote the identification of the causative nucleotide polymorphisms, is to increase the number of crossover events in the mapping population (*e.g.*, Cheng *et al.* 2010). Following Darvasi and Soller (1995) we generated an F_{17} advanced intercross line (AIL) between b3852 and *Sam*, passing the population through additional rounds of recombination – limited to females in *Drosophila* – to expand the genetic map by over 7

fold. We also elected to utilize a selective genotyping approach for the F_{17} population to reduce genotyping costs while maintaining high mapping power (Lander and Botstein 1989; Darvasi and Soller 1992). Because our goal was to fine-map QTL for mPC1, we chose to genotype subsets of the F_{17} individuals with mPC1 values most similar to the parental strains (F_{17} "tail" individuals in Figure 3). These individuals were genotyped for the same 87 SNP markers used for the F_2, with all adjacent markers along a chromosome remaining linked on the expanded F_{17} genetic map.

Figure 4 (bottom panel) presents the results of fine-mapping with IM for both mPC1 and the correlated H:W trait, showing similar results to the F_2 map: The large pericentromeric chromosome 3 mPC1 QTL is preserved on fine-mapping (Q3; LOD = 10.0), a second QTL on 3L (Q2; LOD = 6.8) is present in approximately the same location as the F_2 CIM QTL, and there are QTL on chromosome 2L, including a relatively large QTL in the middle of 2L (Q1; LOD = 6.2). We find no evidence for a QTL in the middle of 3R in the F_{17} IM analysis as we identified with CIM in the F_2, either due to low power to detect it in the F_{17}, or because the F_2 CIM QTL is a artifact (we did not routinely map this QTL when varying the analysis parameters for CIM).

In considering fine-mapping power it should be noted that during laboratory maintenance of the F_{17} population, either drift or selection led to a reduction in *Sam* allele frequency at various points along the genome. One indication of the skewed allele frequency is that the most *Sam*-like F_{17} flies are not as phenotypically extreme as the inbred *Sam* parent (Figure 2), implying a dearth of individuals homozygous for *Sam* alleles at loci contributing to posterior lobe variation. In addition, the frequency of *Sam* alleles is very low along the entire X chromosome, and at the very telomeric end of 3R

in all genotyped F_{17} flies (Figure S4), limiting our power to detect QTL in these regions. We hypothesize the *Sam* genome carries slightly deleterious alleles at several loci, and that individuals homozygous for these alleles were at a competitive disadvantage during creation of the AIL, resulting in a reduction in *Sam* allele frequency. Various multigenerational crossing designs can be used to create AILs while limiting the effects of drift and selection (*e.g.*, Rockman and Kruglyak 2008). Although more cumbersome than typical methods of maintaining large fly populations, such strategies are likely to be beneficial in maintaining a consistent level of mapping power along the genome in AIL-based QTL studies.

To explore our mPC1 QTL data further we used various routines from *R/qtl* (Broman and Sen 2009), beginning with a simple model including the three QTL in which we have the greatest confidence – Q1, Q2, and Q3 – that each have LOD scores > 6 (Figure 4). The 'addqtl' function did not indicate further QTL should be added to the model, although there was some suggestion of an additional QTL at the very tip of 2L. Using a combination of the 'addint' function, which asks whether allowing QTL to interact improves the model fit, and a direct two-dimensional scan for epistatic QTL with 'scantwo', we found that Q1 and Q2 interact. The final model, y = Q1 + Q2 + Q3 + Q1 × Q2, explains an estimated 26.5% of the phenotypic variance (using the 'fitqtl' function). Each QTL contributes 4.6-10.7% to mPC1 variation (Table 3), and the Q1 × Q2 interaction contributes 5.5%. For all three QTL, substitution of a b3852 allele for a *Sam* allele increases mPC1 (giving a more *Sam*-like phenotype), with Q2 and Q3 acting predominantly additively, and Q1 having a large dominance component (Table 3). It seems clear that no single QTL explains a large fraction of the morphological variation

between the parental strains, and that instead trait variation is conferred by the action of a number of relatively small-effect QTL. This is particularly true in light of our somewhat low F_{17} sample size, which has likely resulted in us overestimating QTL effects (Beavis 1994).

The goal of fine-mapping is to reduce QTL map intervals, promoting identification of the causative gene or polymorphism. We succeeded in expanding the map length of the autosomal genome by a factor of >7 between the F_2 and F_{17} generations, and confidence intervals for fine-mapped QTL are smaller than in the course-mapping study. Nonetheless, the three major QTL we identify are still mapped to relatively broad genetic distances (6, 16, and 5 cM for Q1, Q2, and Q3, respectively, on the standard genetic map of *D. melanogaster*), that encompass hundreds of genes (Table 3). Q3 covers a particularly large physical distance since it resides over the centromere of the third chromosome where the rate of recombination is low. We anticipate being able to improve resolution, and decrease the size of the genomic regions implicated, by maintaining the AIL for many additional generations prior to QTL mapping, and by adding markers to increase the number of informative recombination events across QTL intervals (see Macdonald and Long 2007).

Comparing posterior lobe QTL mapped within and between *Drosophila* species

An important challenge in evolutionary genetics is to describe the relationship between intra- and interspecific genetic variation (see Nuzhdin and Reiwitch 2000). Using data from QTL experiments we can ask whether the properties of loci contributing to trait variation within a species are similar to the properties of loci responsible for trait

divergence between species. Such efforts have been used to suggest a shared genetic basis for floral trait variation within *Mimulus guttatus* and between *M. guttatus* and *M. nasutus*, since 11/16 intraspecific QTL map to the same locations in the interspecific cross (Fishman *et al.* 2002; Hall *et al.* 2006). Conversely, due to a lack of overlap between QTL mapped in intra- and interspecies crosses, current evidence suggests there is a qualitative difference in the genetic architecture of courtship song within *D. melanogaster* and between *D. simulans* and *D. sechellia* (Gleason *et al.* 2002; Gleason and Ritchie 2004).

The main result from our study is the identification of at least three moderate-effect QTL contributing to posterior lobe shape between a pair of inbred lines of *D. melanogaster*. The positions of these QTL map to approximately the same locations as QTL mapped in various interspecific crosses (Figure 7 in Liu *et al.* 1996; Figure 3 in Macdonald and Goldstein 1999; Figure 2 in Zeng *et al.* 2000; Figure 6 in Masly *et al.* 2011). Interspecific posterior lobe QTL have also been mapped to the tip of 2L and 3L in these studies, sites where we also find LOD scores just above the QTL significance threshold. This overlap in QTL positions suggests some of the same genes could be responsible for lobe shape variation both within and among species of *Drosophila*.

There is of course a clear caveat: Mapping resolution in all studies considered is relatively low, and with >14,000 genes, just three major chromosomes, and the possibility that a large number of genes influence the trait, these QTL could overlap simply by chance. Short of positionally cloning the causative gene (see Wittkopp *et al.* 2009), progress towards a rigorous comparison of the pattern of genetic variation within and among species is likely to come only once the QTL are resolved to very short

intervals, and can be isolated from the effects of others in introgression lines. In general, large highly-recombinant mapping populations must be employed to achieve this, although in *D. melanogaster* investigators can make use of molecularly-characterized deletions or loss-of-function mutations to implicate putative causative genes via quantitative complementation tests (Long *et al.* 1996; Pasyukova *et al.* 2000).

Candidate gene analysis

In common with the rest of the male and female adult genitalia, the posterior lobe develops from the larval genital imaginal disc. Chatterjee *et al.* (2011) used microarrays to identify 22 euchromatic genes that consistently differ in expression between male and female *D. melanogaster* genital discs across three developmental timepoints (L3 larvae, 6 hours and 20 hours after puparium formation). Seven of these genes were also found by Masly *et al.* (2011), comparing male and female discs in L3 larvae only. In addition, at least two of the genes found in both studies – *Pox neuro* and *Drop* – can be mutated to alter adult posterior lobe morphology (Boll and Noll 2002; Chatterjee *et al.* 2011). We highlight the positions of these 22 loci in Figure 4 (red points above each plot), and note a visually striking overlap between the mPC1 QTL peaks and candidate gene positions, particularly on the autosomes. Interestingly, the classic sex-determining gene *doublesex* (Hildreth 1965) is within the 2-LOD drop for the pericentromeric QTL Q3. This gene was identified as male-biased in the developing *D. melanogaster* genital disc by both Chatterjee *et al*. (2011) and Masly *et al.* (2011), and was also one of the genes the latter work identified as differentially expressed between *D. mauritiana* and *D. sechellia* in male genital discs.

To test for a statistical association between our QTL results and these candidate loci we used a resampling procedure (Keightley *et al.* 1998): One million sets of 18 autosomal loci were randomly sampled, and the mPC1 LOD scores at the 18 positions summed (the X chromosome was ignored because the low *Sam* allele frequency on this chromosome in the F_{17} likely compromised mapping power). This gives a distribution of the expected LOD scores assuming no relationship between our phenotype and the Chatterjee *et al.* (2011) candidate genes. The sum of the PC1 LOD scores at the actual locations of the 18 autosomal candidate genes is 68.7, which is in the top 1% of the null distribution (mean = 39.1, standard deviation = 9.33), indicating a significant association between the two datasets. Thus, those genes that show sex-biased gene expression in genital discs, and are present within QTL intervals (see legend to Figure 4), are plausible candidates to harbor natural genetic variants contributing to posterior lobe shape.

ACKNOWLEDGEMENTS

We thank Charles Nicolet at the UC Davis Genome Center for help and advice with SNP genotyping, Steven Hoofer for assistance with the dissections, and Jennifer Gleason, John Kelly and two anonymous reviewers for valuable comments on the manuscript. This work was supported by a National Institutes of Health Grant R01 GM085260 to SJM.

LITERATURE CITED

Arnqvist, G., and I. Danielsson, 1999 Copulatory behavior, genital morphology, and male fertilization success in water striders. Evolution **53:** 147–156.

Ashburner, M., K. G. Golic, and R. S. Hawley, 2005 *Drosophila: A laboratory handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Beavis, W. D., 1994 The power and deceit of QTL experiments: lessons from comparative QTL studies. Proceedings of the 49th Annual Corn & Sorghum Industry Research Conference. American Seed Trade Association, Washington, DC, pp. 250–266.

Boll, W., and M. Noll, 2002 The *Drosophila Pox neuro* gene: control of male courtship behavior and fertility as revealed by a complete dissection of all enhancers. Development **129:** 5667–5681.

Broman, K. W., and S. Sen, 2009 *A Guide to QTL Mapping with R/qtl*. Springer, New York.

Broman, K. W., S. Sen, S. E. Owens, A. Manichaikul, E. M. Southard-Smith *et al.*, 2006 The X chromosome in quantitative trait locus mapping. Genetics **174:** 2151–2158.

Chatterjee, S. S., L. D. Uppendahl, M. A. Chowdhury, P-L. Ip, and M. L. Siegal, 2011 The female-specific Doublesex isoform regulates pleiotropic transcription factors to pattern genital development in *Drosophila*. Development **138:** 1099–1109.

Cheng, R., J. E. Lim, K. E. Samocha, G. Sokoloff, M. Abney, A. D. Skol, and A. A. Palmer, 2010 Genome-wide association studies and the problem of relatedness

among advanced intercross lines and other highly recombinant populations. Genetics **185:** 1033–1044.

Churchill, G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. Genetics **138:** 963–971.

Clark, A. G., M. B. Eisen, D. R. Smith, C. M. Bergman, B. Oliver *et al.*, 2007 Evolution of genes and genomes on the *Drosophila* phylogeny. Nature **450:** 203–218.

Cornforth, T. W., and A. D. Long, 2003 Inferences regarding the numbers and locations of QTLs under multiple-QTL models using interval mapping and composite interval mapping. Genet. Res. **82:** 139–149.

Coyne, J. A., 1983 Genetic basis of differences in genital morphology among three sibling species of *Drosophila*. Evolution **37:** 1101–1118.

Coyne, J. A., and M. Kreitman, 1986 Evolutionary genetics of two sibling species, *Drosophila simulans* and *D. sechellia*. Evolution **40:** 673–691.

Danielsson, I., and C. Askenmo, 1999 Male genital traits and mating interval affect male fertilization success in the water strider *Gerris lacustris*. Behav. Ecol. Sociobiol. **46:** 149–156.

Darvasi, A., and M. Soller, 1992 Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theor. Appl. Genet. **85:** 353–359.

Darvasi, A., and M. Soller, 1995 Advanced intercross lines, an experimental population for fine genetic mapping. Genetics **141:** 1199–1207.

Eberhard, W. G., 1985 *Sexual Selection and Animal Genitalia*. Harvard University Press, Cambridge, MA.

Eberhard, W. G., 2010 Evolution of genitalia: theories, evidence, and new directions. Genetica **138:** 5–18.

Eberhard, W. G., 2011 Experiments with genitalia: a commentary. TREE **26:** 17–21.

Eberhard, W. G., B. A. Huber, R. L. Rodriguez, S. R. D. Briceno, I. Salas *et al.*, 1998 One size fits all? Relationships between the size and degree of variation in genitalia and other body parts in twenty species of insects and spiders. Evolution **52:** 415–431.

Ferson, S., F. J. Rohlf and R. K. Koehn, 1985 Measuring shape variation of two-dimensional outlines. Syst. Zool. **34:** 59–68.

Fishman, L., A. J. Kelly, and J. H. Willis, 2002 Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. Evolution **56:** 2138–2155.

Flint, J., and T. F. Mackay, 2009 Genetic architecture of quantitative traits in mice, flies, and humans. Genome Res. **19:** 723–733.

Gleason, J. M., and M. G. Ritchie, 2004 Do quantitative trait loci (QTL) for a courtship song difference between *Drosophila simulans* and *D. sechellia* coincide with candidate genes and intraspecific QTL? Genetics **166:** 1303–1311.

Gleason, J. M., S. V. Nuzhdin, and M. G. Ritchie, 2002 Quantitative trait loci affecting a courtship signal in *Drosophila melanogaster*. Heredity **89:** 1–6.

Hall, M. C., C. J. Basten, and J. H. Willis, 2006 Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. Genetics **172:** 1829–1844.

Hildreth, P. E., 1965 *Doublesex*, a recessive gene that transforms both males and females of *Drosophila* into intersexes. Genetics **51:** 659–678.

Hoskey, D. J., and P. Stockley, 2004 Sexual selection and genital evolution. TREE **19:** 87–93.

Jagadeeshan, S., and R.S. Singh, 2006 A time-sequence functional analysis of mating behaviour and genital coupling in *Drosophila*: role of cryptic female choice and male sex-drive in the evolution of genitalia. J. Evol. Biol. **19:** 1058–1070.

Keightley, P. D., K. H. Morris, A. Ishikawa, V. M. Falconer, and F. Oliver, 1998 Test of candidate gene-quantitative trait locus association applied to fatness in mice. Heredity **81:** 630–637.

Kuhl, F. P., and C. R. Giardina, 1982 Elliptic Fourier features of a closed contour. Comput. Graphics Image Processing **18:** 236–258.

Lander, E. S., and D. Botstein, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

Laurie, C. C., J. R. True, J. Liu, and J. M. Mercer, 1997 An introgression analysis of quantitative trait loci that contribute to a morphological difference between *Drosophila simulans* and *D. mauritiana*. Genetics **145:** 339–348.

Liu, J., J. M. Mercer, L. F. Stam, G. C. Gibson, Z-B. Zeng *et al.*, 1996 Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. Genetics **145:** 1129–1145.

Long, A. D., S. L. Mullaney, T. F. C. Mackay, and C. H. Langley, 1996 Genetic interactions between naturally occurring alleles at Quantitative Trait Loci and

mutant alleles at candidate loci affecting bristle number in *Drosophila melanogaster*. Genetics **144:** 1497–1510.

Lyman, R. F., F. Lawrence, S. V. Nuzhdin, and T. F. C. Mackay, 1996 Effects of single *P*-element insertions on bristle number and viability in *Drosophila melanogaster*. Genetics **143:** 277–292.

Macdonald, S. J., and D. B. Goldstein, 1999 A quantitative genetic analysis of male sexual traits distinguishing the sibling species *Drosophila simulans* and *D. sechellia*. Genetics **153:** 1683–1699.

Macdonald, S. J., and A. D. Long, 2007 Joint estimates of Quantitative Trait Locus effect and frequency using synthetic recombinant populations of *Drosophila melanogaster*. Genetics **176:** 1261–1281.

Macdonald, S. J., T. Pastinen, A. Genissel, T. W. Cornforth and A. D. Long, 2005 A low-cost open-source SNP genotyping platform for association mapping applications. Genome Biol. **6:** R105.

Mackay, T. F., 2010 Mutations and quantitative genetic variation: lessons from *Drosophila*. Philos. Trans. R. Soc. B-Biol. Sci. **365:** 1229–1239.

Manichaikul, A., A. A. Palmer, S. Sen, and K. W. Broman, 2007 Significance thresholds for quantitative trait locus mapping under selective genotyping. Genetics **177:** 1963–1966.

Masly, J. P., J. E. Dalton, S. Srivastava, L. Chen, and M. N. Arbeitman, 2011 The genetic basis of rapidly evolving male genital morphology in *Drosophila*. Genetics *in press*.

Mitchell-Olds, T., J. H. Willis, and D. B. Goldstein, 2007 Which evolutionary processes influence natural genetic variation for phenotypic traits? Nature Rev. Genet. **8:** 845–856.

Nuzhdin, S. V., and S. G. Reiwitch, 2000 Are the same genes responsible for intra- and interspecific variability for sex comb tooth number in *Drosophila*? Heredity **84:** 97–102.

Orr, H. A., 1998 Testing natural selection *vs.* genetic drift in phenotypic evolution using quantitative trait locus data. Genetics **149:** 2099–2104.

Pasyukova, E. G., C. Vieira, and T. F. C. Mackay, 2000 Deficiency mapping of Quantitative Trait Loci affecting longevity in *Drosophila melanogaster*. Genetics **156:** 1129–1146.

Robertson, H. M., 1988 Mating asymmetries and phylogeny in the *Drosophila melanogaster* species complex. Pac. Sci. **42:** 72–80.

Rockman, M. V., and L. Kruglyak, 2008 Breeding designs for recombinant inbred advanced intercross lines. Genetics **179:** 1069–1078.

Sasabe, M., Y. Takami, and T. Sota, 2010 QTL for the species-specific male and female genital morphologies in *Ohomopterus* ground beetles. Mol. Ecol. 19: 5231–5239.

Schafer, M. A., J. Routtu, J. Vieira, A. Hoikkala, M. G. Ritchie, and C. Schlotterer, 2011 Multiple quantitative trait loci influence intra-specific variation in genital morphology between phylogenetically distinct lines of *Drosophila montana*. J. Evol. Biol. *in press*.

Sen, S., and G. A. Churchill, 2001 A statistical framework for quantitative trait mapping. Genetics **159:** 371–387.

Sen, S., J. M. Satagopan, and G. A. Churchill, 2005 Quantitative trait locus study design from an information perspective. Genetics **170:** 447–464.

Shingleton, A. W., C. M. Estep, M. V. Driscoll, and I. Dworkin, 2009 Many ways to be small: different environmental regulators of size generate distinct scaling relationships in *Drosophila melanogaster*. Proc. Biol. Sci. **276:** 2625–2633.

Templeton, A. R., 1981 Mechanisms of speciation - A population genetic approach. Ann. Rev. Evol. Syst. **12:** 23–48.

Wittkopp, P. J., E. E. Stewart, L. L. Arnold, A. H. Neidert, B. K. Haerum *et al.*, 2009 Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in *Drosophila*. Science **326:** 540–544.

Zeng, Z-B., 1994 Precision mapping of quantitative trait loci. Genetics **136:** 1457–1468.

Zeng, Z. B., J. Liu, L. F. Stam, C. H. Kao, J. M. Mercer *et al.*, 2000 Genetic architecture of a morphological shape difference between two *Drosophila* species. Genetics **154:** 299–310.

SUPPLEMENTARY DATA FILES

File S1 Development of SNP markers. The nucleotide sequence surrounding each SNP is provided, along with the position of the SNP in the *D. melanogaster* reference genome (release 5.2), and the GoldenGate assay genotyping score provided by Illumina.

File S2 Raw phenotypes and genotypes for all F_2 individuals. All phenotypes discussed in the text are provided for each individual, along with their genotypes (A = homozygous b3852 genotype, H = heterozygote, B = homozygous *Sam* genotype). Genetic positions estimated via *R/qtl* are also provided for all markers. This file is suitable for direct input into *R/qtl*.

File S3 Raw phenotypes and genotypes for all F_{17} individuals. See legend for File S2.

Table 1 Posterior lobe morphology shows variation across *D. melanogaster* strains

Strain [a]	*N* [b]	Lobe area ($\times 10^{-3}$ mm^2)	Lobe height ($\times 10^{-3}$ mm)	Lobe width ($\times 10^{-3}$ mm)	H:W [c]	sPC1 ($\times 10^{-4}$) [d]	sPC2 ($\times 10^{-4}$) [d]	sPC3 ($\times 10^{-4}$) [d]
b1	24	3.43 (0.228)	56.0 (2.86)	59.0 (1.92)	0.95 (0.044)	37.2 (13.73)	–26.8 (15.15)	23.7 (17.29)
b3839	32	3.62 (0.174)	56.9 (2.22)	61.4 (1.92)	0.93 (0.046)	–7.7 (16.01)	–15.2 (15.06)	–4.5 (17.04)
b3841	35	3.32 (0.235)	51.7 (2.56)	62.4 (2.77)	0.83 (0.051)	–3.4 (20.82)	8.5 (13.00)	–6.6 (18.08)
b3844	36	3.17 (0.168)	51.1 (2.01)	60.5 (1.91)	0.84 (0.039)	16.3 (14.44)	13.5 (12.62)	10.7 (14.03)
b3846	25	3.67 (0.160)	51.5 (1.46)	70.0 (2.83)	0.74 (0.040)	31.9 (20.43)	8.7 (12.63)	–31.8 (13.41)
b3852	14	3.14 (0.153)	55.7 (2.40)	51.8 (1.54)	1.08 (0.046)	–94.3 (12.15)	12.1 (10.74)	–1.3 (14.89)
b3864	15	3.95 (0.170)	57.4 (2.20)	66.3 (2.26)	0.87 (0.051)	23.9 (16.62)	–33.8 (15.90)	–25.7 (15.10)
b3870	23	3.22 (0.155)	51.5 (1.81)	62.8 (1.56)	0.82 (0.029)	–7.9 (13.21)	22.0 (10.98)	1.7 (12.32)
b3875	14	3.64 (0.197)	58.1 (2.10)	62.2 (2.30)	0.94 (0.042)	–5.9 (15.99)	–14.1 (12.11)	–8.0 (12.36)
b3886	18	3.47 (0.242)	54.2 (2.94)	63.0 (2.23)	0.86 (0.054)	–14.1 (17.84)	8.6 (16.55)	–19.8 (13.97)
Samarkand ry^{506}	13	3.06 (0.247)	46.0 (1.95)	66.2 (3.55)	0.70 (0.033)	62.8 (16.32)	34.9 (11.74)	18.1 (18.37)
t14021-0231.0	18	3.53 (0.194)	58.3 (2.34)	57.9 (2.30)	1.01 (0.050)	3.1 (18.63)	–26.2 (14.96)	8.4 (12.83)
t14021-0231.1	12	3.25 (0.126)	50.3 (1.83)	62.7 (1.42)	0.80 (0.035)	11.6 (11.63)	18.9 (10.66)	–4.8 (11.25)
t14021-0231.4	21	3.29 (0.250)	50.4 (2.87)	65.6 (2.35)	0.77 (0.045)	34.4 (24.00)	6.8 (11.92)	9.8 (19.07)
t14021-0231.7	42	3.21 (0.222)	54.7 (2.86)	55.0 (1.78)	1.00 (0.049)	–51.6 (16.87)	–8.2 (20.73)	12.2 (12.67)

Values represent the strain mean (standard deviation) for various measures of posterior lobe size and shape.

[a] Names of the 15 strains used. b3852 and *Samarkand ry*^{506} were used as parents for the F_2 QTL mapping study.

[b] Number of individuals phenotyped per strain. For each fly the phenotype data from just one lobe was used.

[c] H:W is the ratio of lobe height to lobe width.

[d] The first three principal components from the species diversity PCA employing 100 elliptic Fourier coefficients per lobe. These represent orthogonal aspects of posterior lobe shape, and explain 53.0% (sPC1), 18.8% (sPC2), and 16.3% (sPC3) of shape variation across the 15 strains. All other PCs explain less than 5% of lobe variation.

Table 2 Correlations among lobe phenotypes in the two mapping panels

	Area	Height	Width	H:W [a]	mPC1 [b]	mPC2 [b]	mPC3 [b]
Area	—	0.63***	0.49***	–0.04	0.17	0.05	0.95***
Height	0.76***	—	–0.31	0.73***	–0.57**	0.26	0.70***
Width	0.56***	–0.04	—	–0.87***	0.81***	–0.39***	0.33***
H:W [a]	0.12	0.71***	–0.73***	—	–0.86***	0.42***	0.10
mPC1 [b]	0.16	–0.35**	0.61***	–0.67***	—	NA	NA
mPC2 [b]	0.04	0.20	–0.42***	0.43***	NA	—	NA
mPC3 [b]	0.95***	0.70***	0.53***	0.10	NA	NA	—

Correlations between traits in the F_2 are above the diagonal (N = 367), and correlations in the F_{17} are below the diagonal (N = 344). Only a single lobe was used from each individual. Asterisks are used to represent significance level (* = 1×10^{-5}, ** = 1×10^{-10}, *** = 1×10^{-15}).

[a] H:W is the ratio between lobe height and lobe width.

[b] The three major principal components explain 62.6% (mPC1), 17.1% (mPC2), and 12.2% (mPC3) of the shape variation in the mapping experiment PCA. Principal components are orthogonal, so correlations among them using the full dataset will be zero by design (and NA values are presented). Although this is not true when considering only a subset of the individuals used in a PCA, there are no significant correlations among principal components in either the F_2 or the F_{17} (data not shown).

Table 3 Details of the fine-mapped mPC1 shape QTL

	Q1	Q2	Q3
Chromosome	2L	3L	3c [a]
Peak LOD	6.2	6.8	10.0
Variance explained (%)	10.7	8.2	4.6
Additive effect ($\times 10^{-4}$)	4.51	6.02	9.61
Dominance effect ($\times 10^{-4}$)	6.76	1.07	1.27
Interval (cM, expanded scale)	75 - 94	78 - 128	186 - 200
Interval (cM, regular scale)	25 - 31	22 - 38	45 - 50
Cytology	27E - 29A	66B - 69B	75F - 86C
Physical size (Mb)	1.17	4.19	11.95
Number of genes	147 (7)	555 (19)	1,383 (69)

QTL intervals are based on a 2-LOD drop from each peak on the expanded F_{17} genetic map. The genetic intervals on the regular, unexpanded genetic map, as well as the cytological intervals were inferred from the expanded genetic map by using the known positions of markers, and physical-to-genetic distance conversion tables on FlyBase. The physical size of each QTL interval, and the number of protein-coding genes (noncoding RNA genes) are also given.

[a] Implicated QTL interval spans the chromosome 3 centromere.

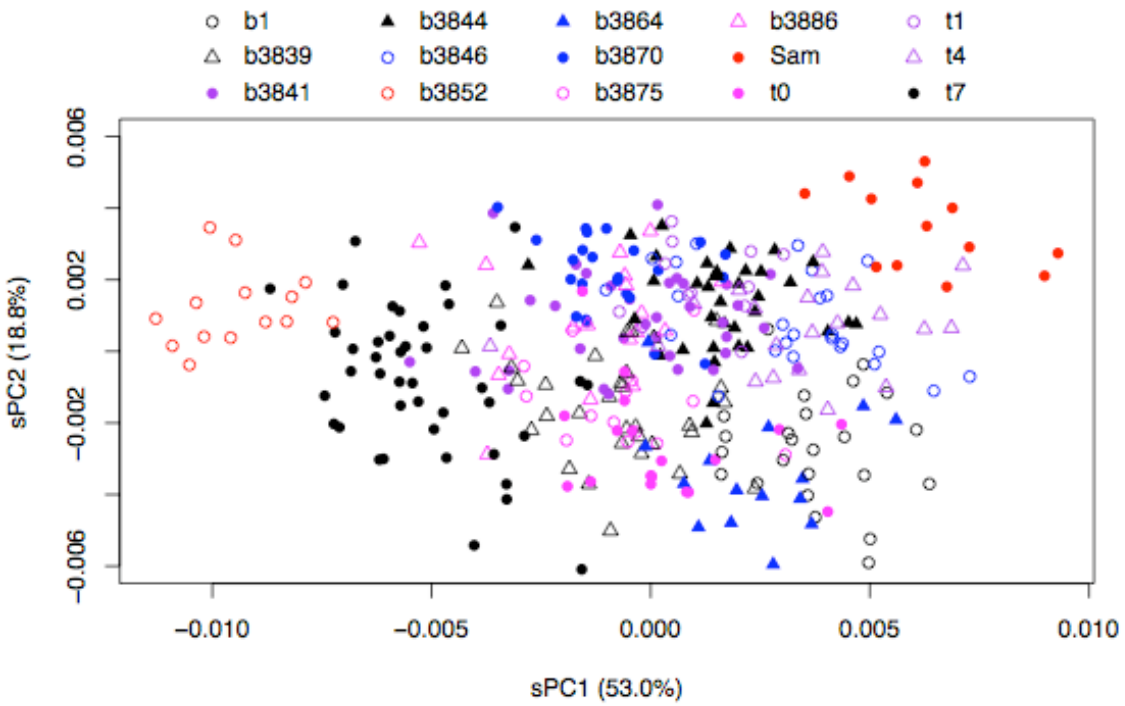


Figure 1 Variation in posterior lobe shape in *D. melanogaster*. Lobe outlines from a number of males (one lobe per individual) from 15 inbred lines were subjected to elliptic Fourier analysis, and the resulting coefficients used in a principal components analysis (the species diversity PCA). Considerable variation in shape among strains is shown for the two major principal components, sPC1 and sPC2. Strains in red symbols are those chosen as the parents for QTL mapping.

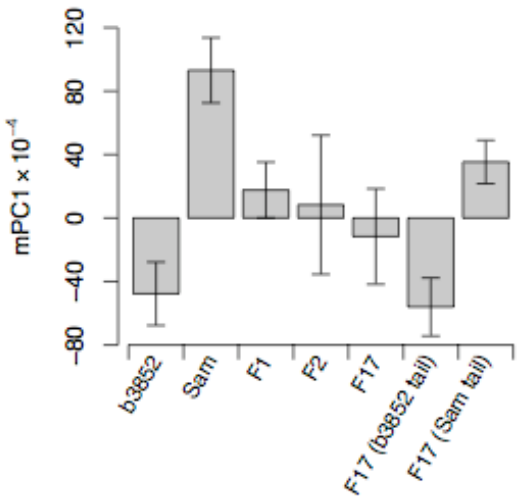


Figure 2 mPC1 lobe shape variation in b3852, *Sam*, F_1, and recombinant genotype classes. Each bar shows the mean (± SD) of multiple individuals, taking just a single lobe from each fly: b3852 (*N* = 29), *Sam* (*N* = 25), F_1 (*N* = 21), F_2 (*N* = 367), F_{17} (*N* = 344), F_{17} b3852 tail (*N* = 47), and F_{17} Sam tail (*N* = 47). F_1 males derived from reciprocal parental crosses have similar shapes and were averaged here. The groups of F_{17} "tail" flies are the most extreme individuals from either tail of the F_{17} phenotypic distribution.

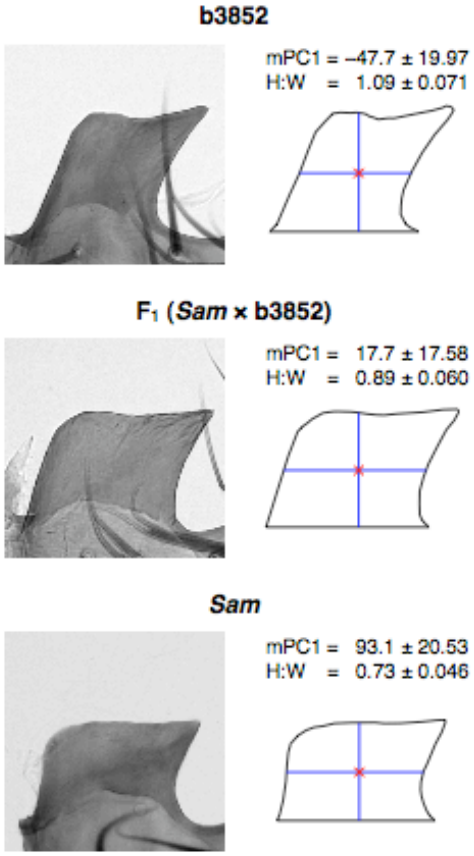


Figure 3 Morphology of the posterior lobe in the progenitors of the QTL mapping panels. On the left a representative lobe image is presented for the two parental strains and the F_1 hybrid (the result of a *Sam* female × b3852 male cross). The closed outlines derived from these images used for shape/size analysis are presented on the right. The red cross within each outline is the centroid, and blue lines represent lobe height and width. The mean (± SD) of the mPC1 ($\times 10^{-4}$) and height : width ratio (H:W) shape measures for each genotype highlight the inverse correlation between these two measures. Phenotype means are calculated from a single lobe from 29 b3852, 21 F_1, and 25 *Sam* males.

Figure 4

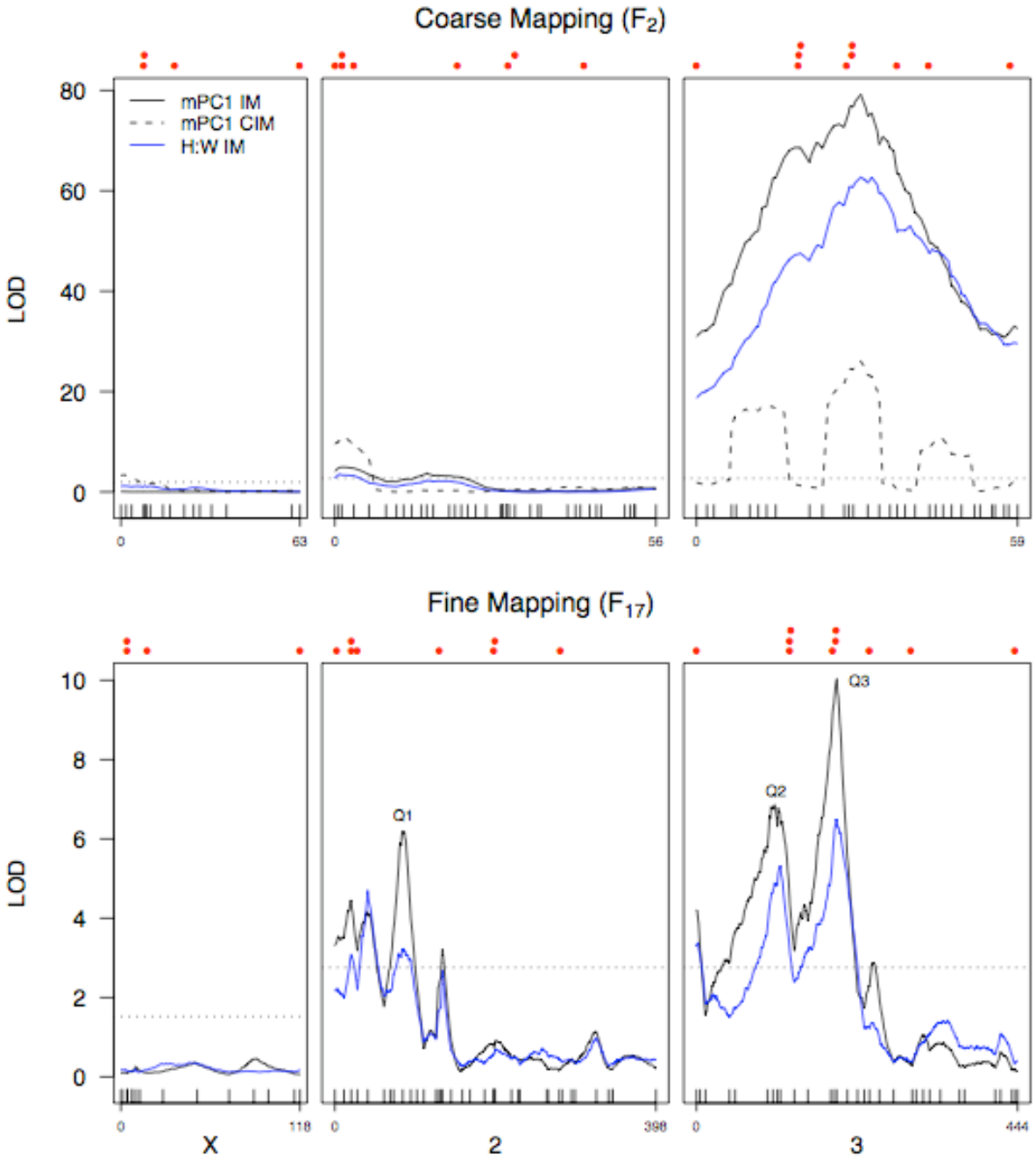

Coarse Mapping (F_2)
mPC1 IM
mPC1 CIM
H:W IM
LOD
80
60
40
20
0
0
63
0
56
0
59
Fine Mapping (F_{17})
LOD
10
8
6
4
2
0
Q1
Q2
Q3
0
118
X
0
398
2
0
444
3

Figure 4 Autosomal QTL contribute to lobe shape variation between b3852 and *Sam*. Likelihood profiles from interval mapping (IM) are shown for mPC1 and the height : width ratio (H:W) for both the coarse- and fine-mapping experiments. In addition, the profile from composite interval mapping (CIM) is shown for mPC1 in the coarse-mapping experiment. The three major fine-mapped mPC1 QTL discussed in the text (Q1, Q2, Q3) are highlighted. The horizontal dotted line represents a 5% significance threshold, and since thresholds for each trait were very similar we conservatively present only the highest threshold. The same set of 87 SNP markers was used for both mapping experiments (ticks along the *x*-axis), but care should be taken when comparing the two sets of plots as map lengths differ (given in cM, F_{17} length > F_2 length), and relative marker spacing is not necessarily preserved. Above the likelihood profiles we mark with solid red circles the positions of 22 plausible candidate genes that show sex-biased gene expression in genital discs (Chatterjee *et al.* 2011). These are (from left to right): X chromosome = *CG4766*, *Nep1*, *lz*, and *FucTC*; chromosome 2 = *al*, *CG4267*, *CG31686*, *msl-2*, *salr*, *ap*, *Wnt2*, and *Poxn*; chromosome 3 = *bab1*, *toe*, *eyg*, *caup*, *AP-2*, *dsx*, *Ctr1B*, *abd-A*, *bnl*, and *Dr*.

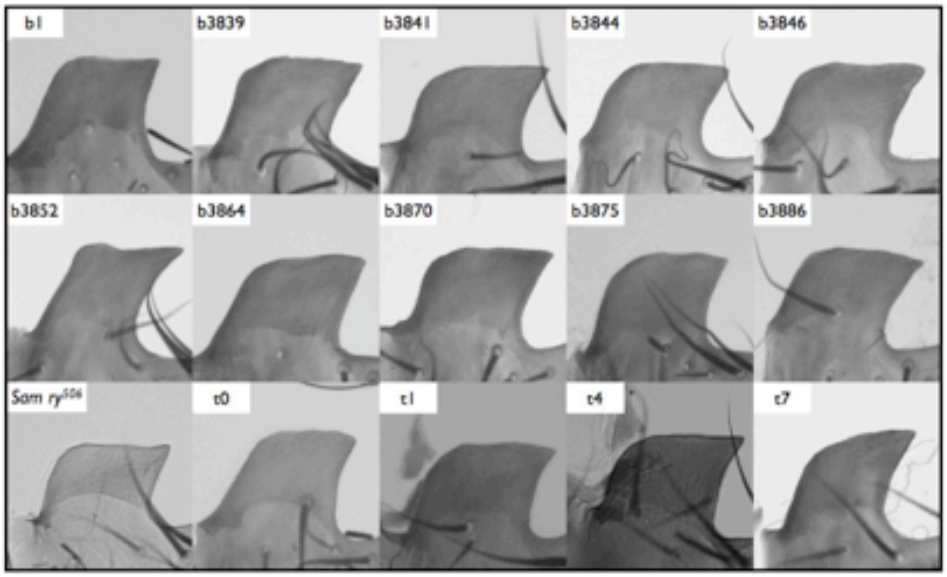


Figure S1 Posterior lobes from 15 *D. melanogaster* inbred lines. A single, representative lobe is shown for each of the strains used in the study: b1, b3839, b3841, b3844, b3846, b3852 (mapping strain), b3864, b3870, b3875, b3886, *Sam* ry^{506} (mapping strain), t0, t1, t4, and t7. All images were taken at the same magnification.

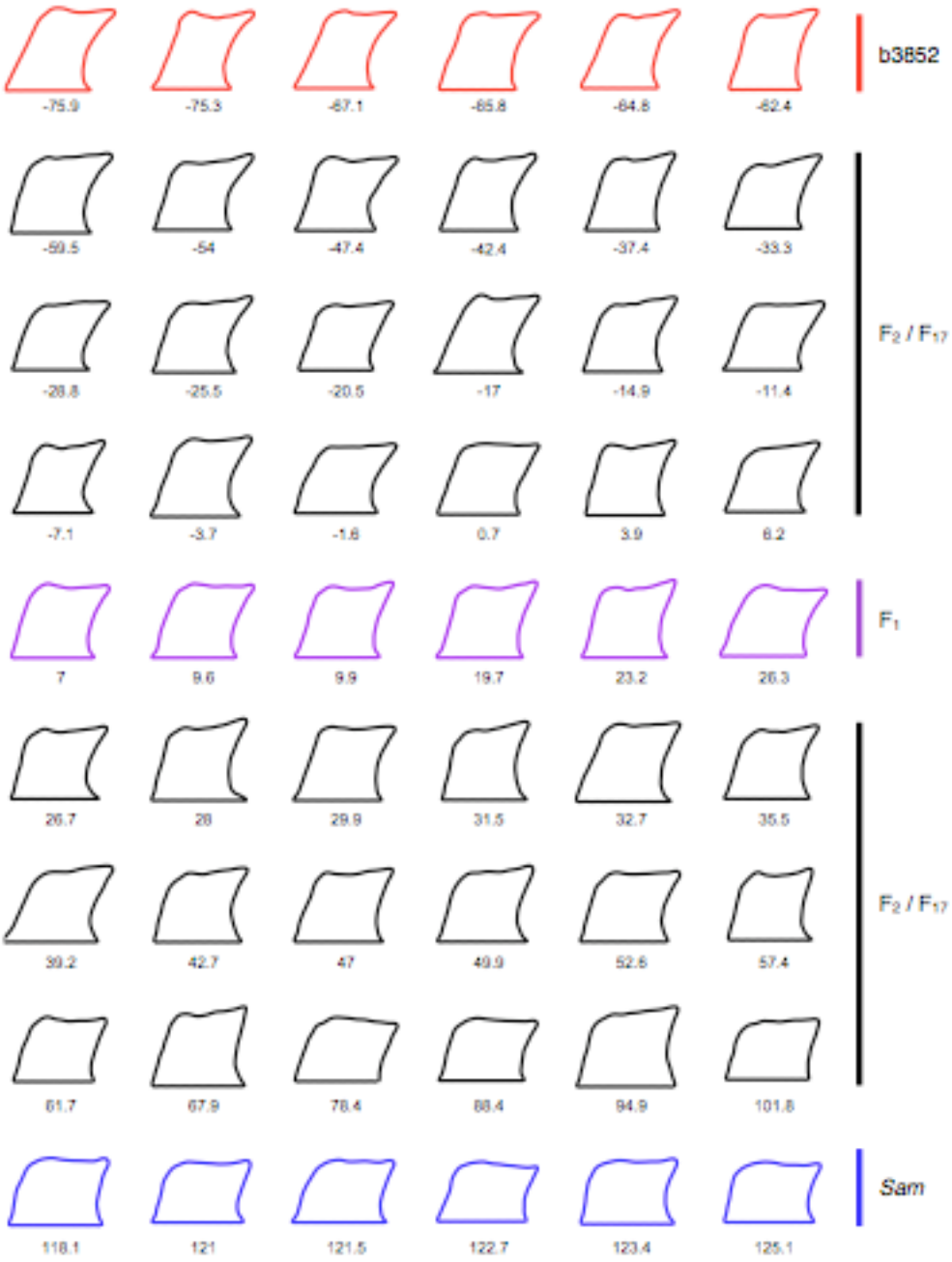
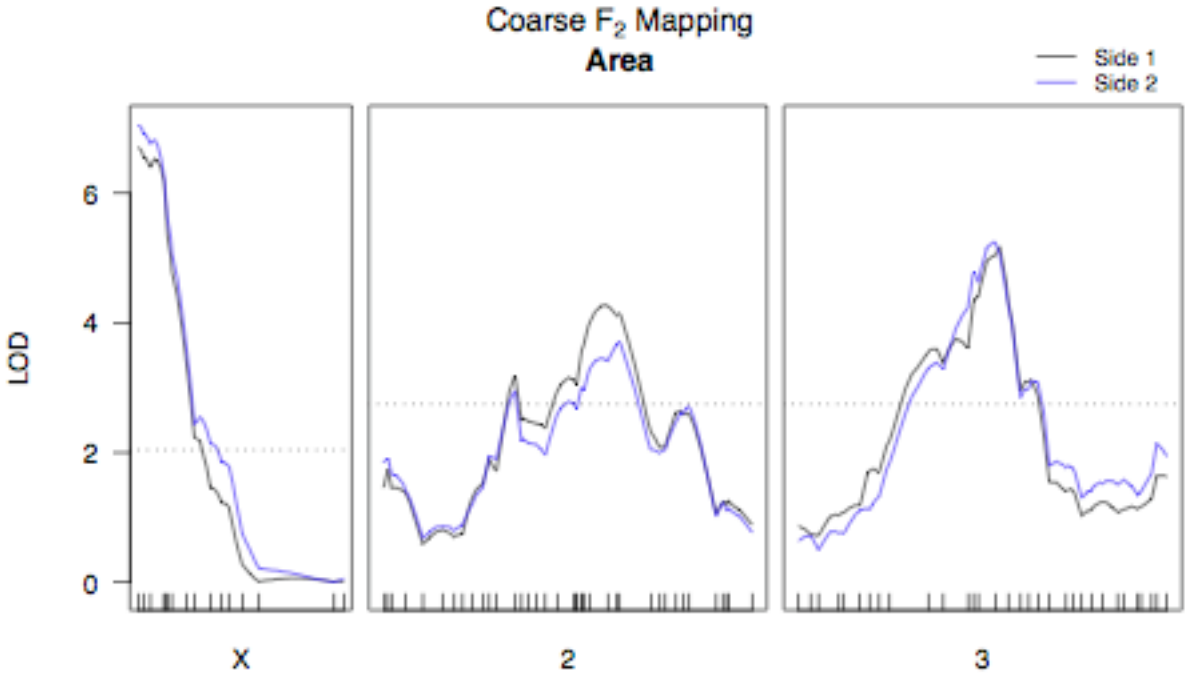


Figure S2 Outlines of lobes from different genotypes showing the change in shape associated with the mPC1 measure. Lobes from a subset of b3852 (red), *Sam* (blue), F_1 (purple), and recombinant F_2 or F_{17} individuals (black) are shown, sorted by their mPC1 score ($\times 10^{-4}$).

Figure S3


Coarse F2 Mapping
Area
Side 1
Side 2
LOD
6
4
2
0
X
2
3

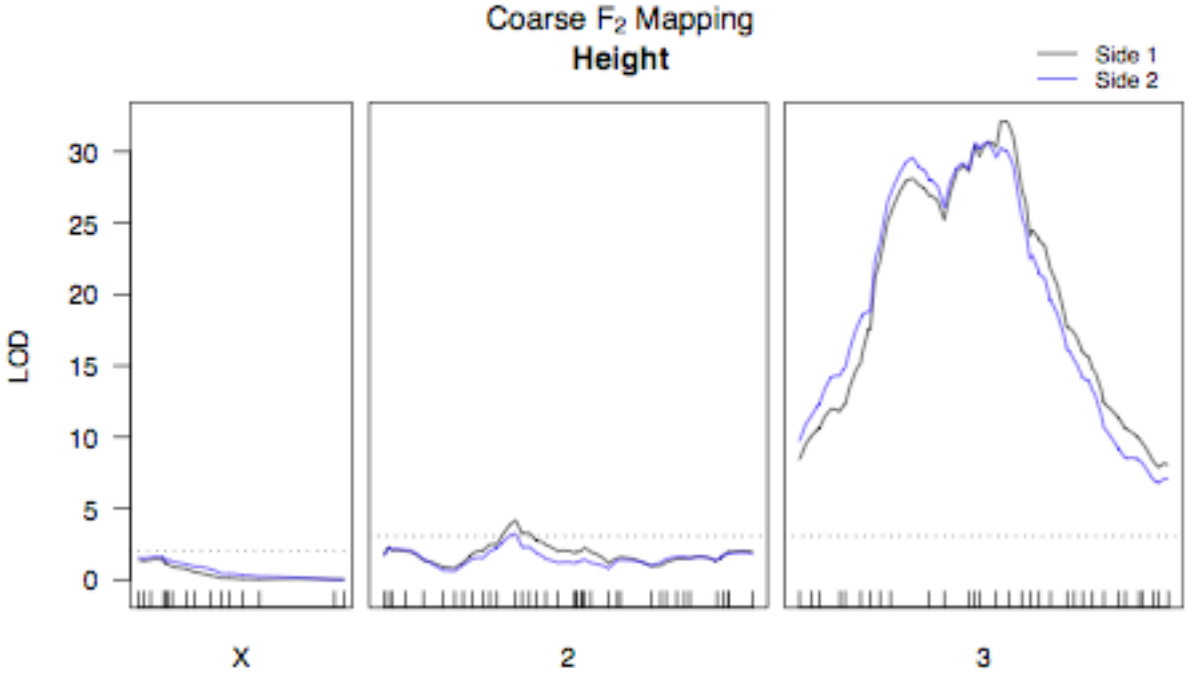
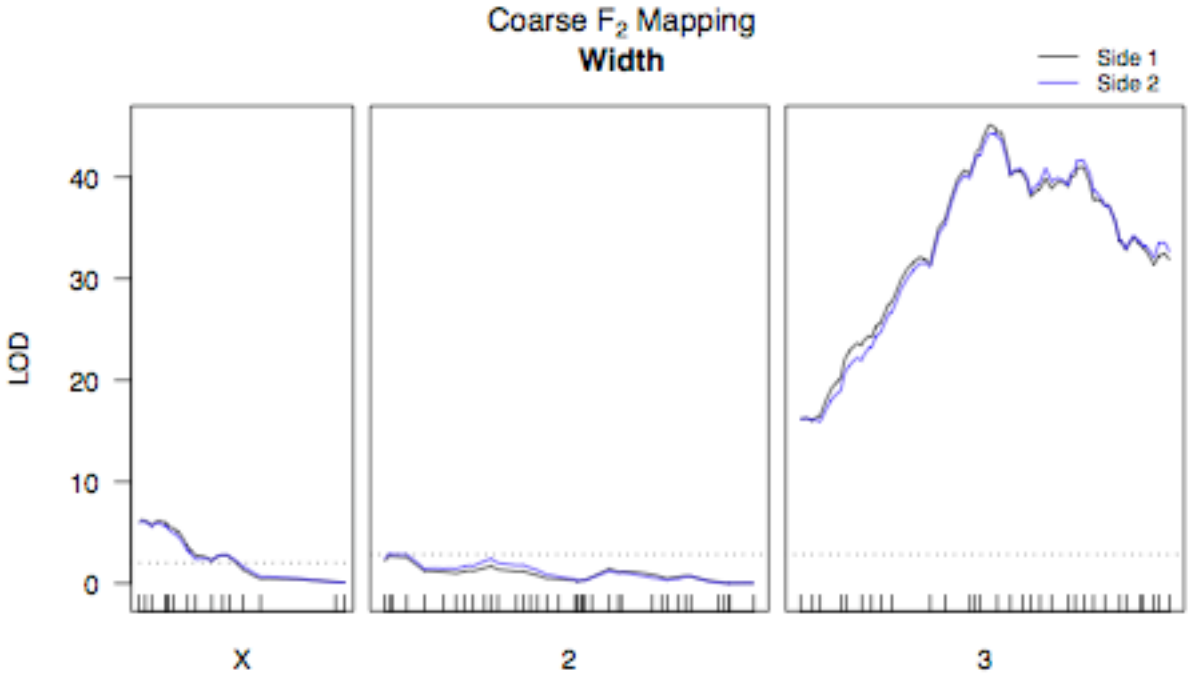

Coarse F2 Mapping
Height
Side 1
Side 2
LOD
30
25
20
15
10
5
0
X
2
3

Figure S3


Coarse F2 Mapping
Width
Side 1
Side 2
LOD
40
30
20
10
0
X
2
3

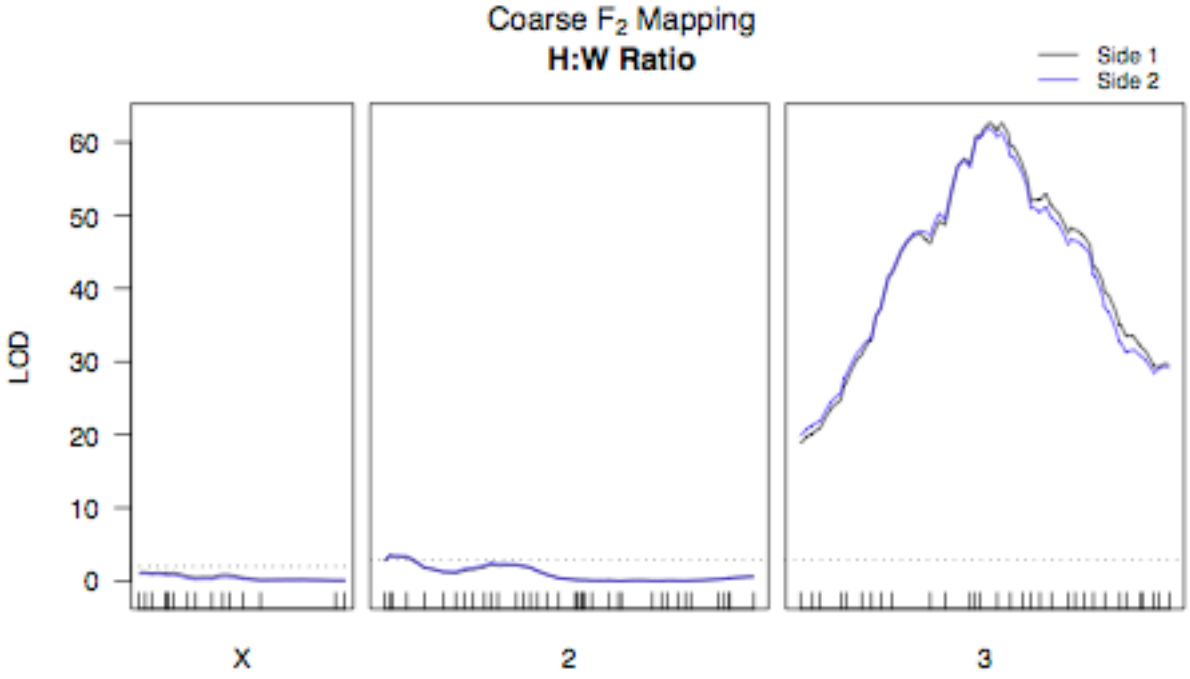
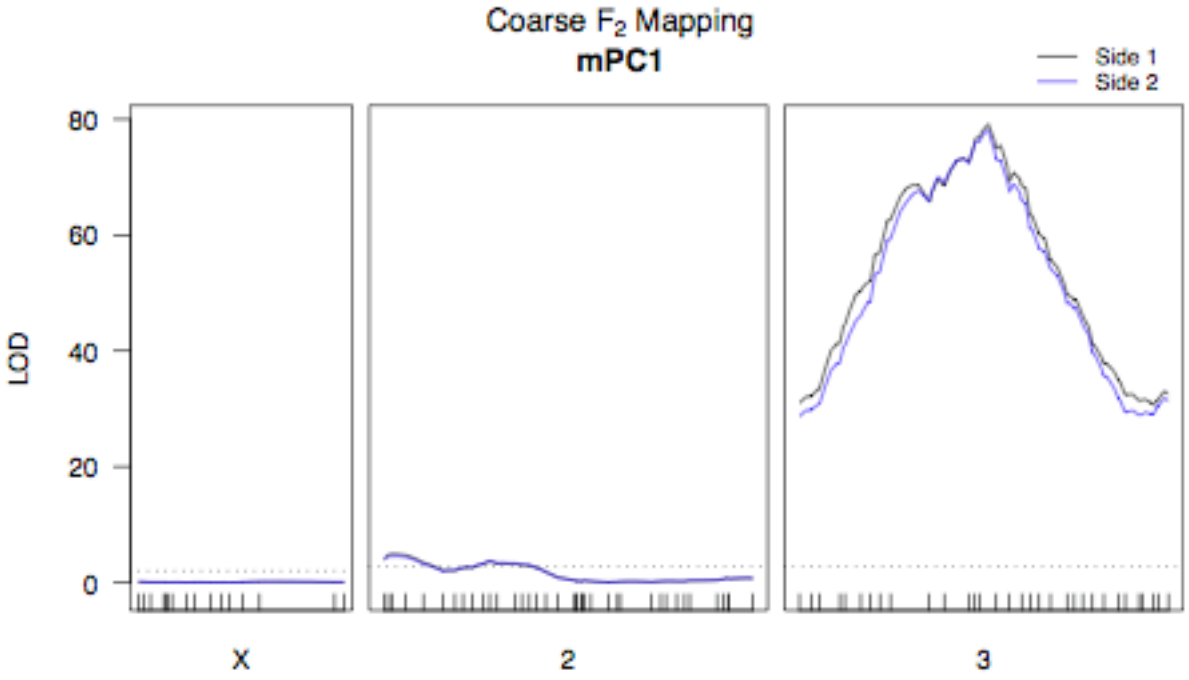

Coarse F2 Mapping
H:W Ratio
Side 1
Side 2
LOD
60
50
40
30
20
10
0
X
2
3

Figure S3


Coarse F2 Mapping
mPC1
Side 1
Side 2
LOD
80
60
40
20
0
X
2
3

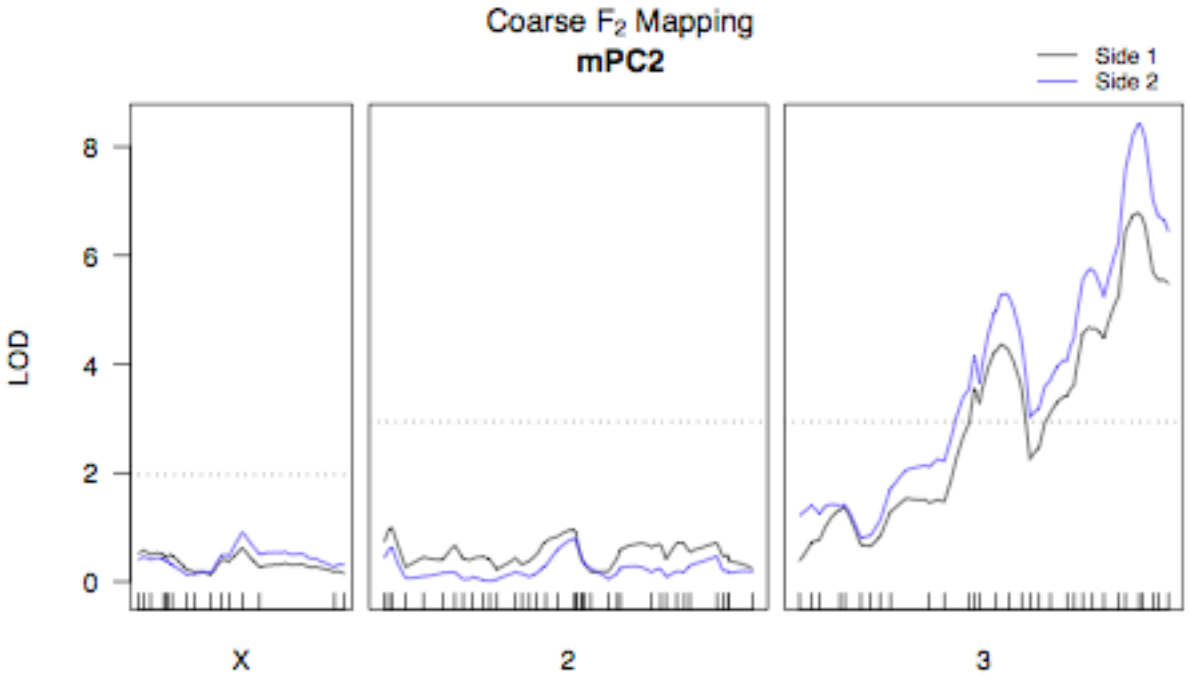
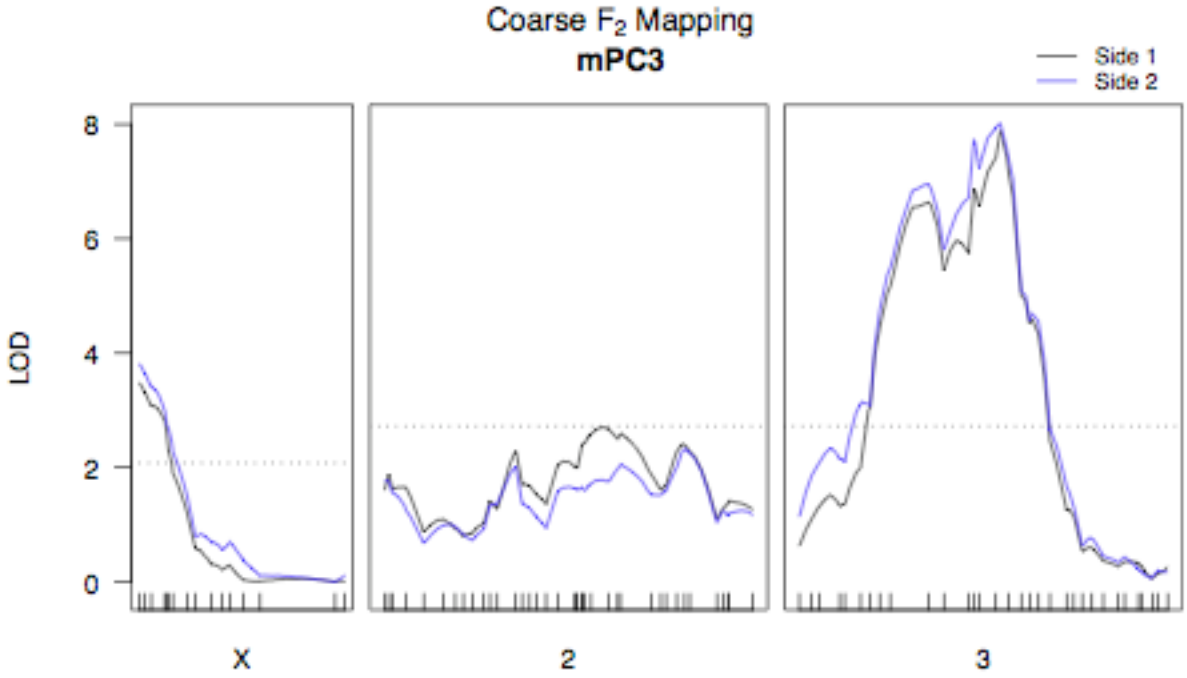

Coarse F2 Mapping
mPC2
Side 1
Side 2
LOD
8
6
4
2
0
X
2
3

Figure S3


Coarse F2 Mapping
mPC3
Side 1
Side 2
LOD
8
6
4
2
0
X
2
3

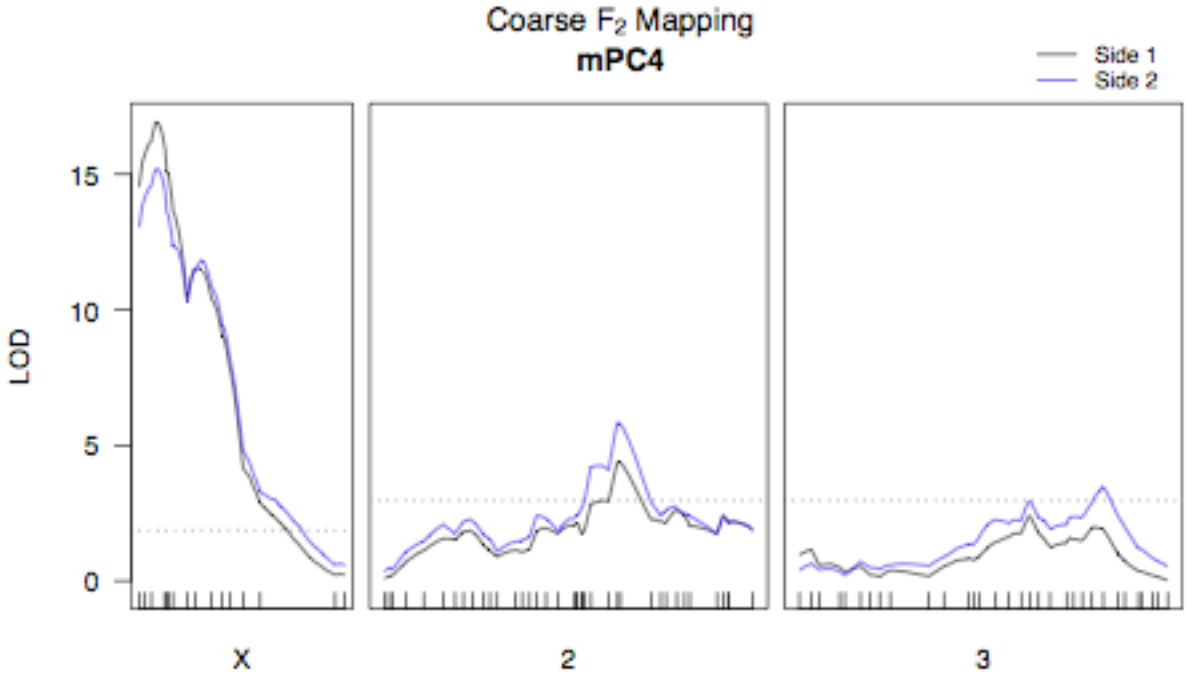
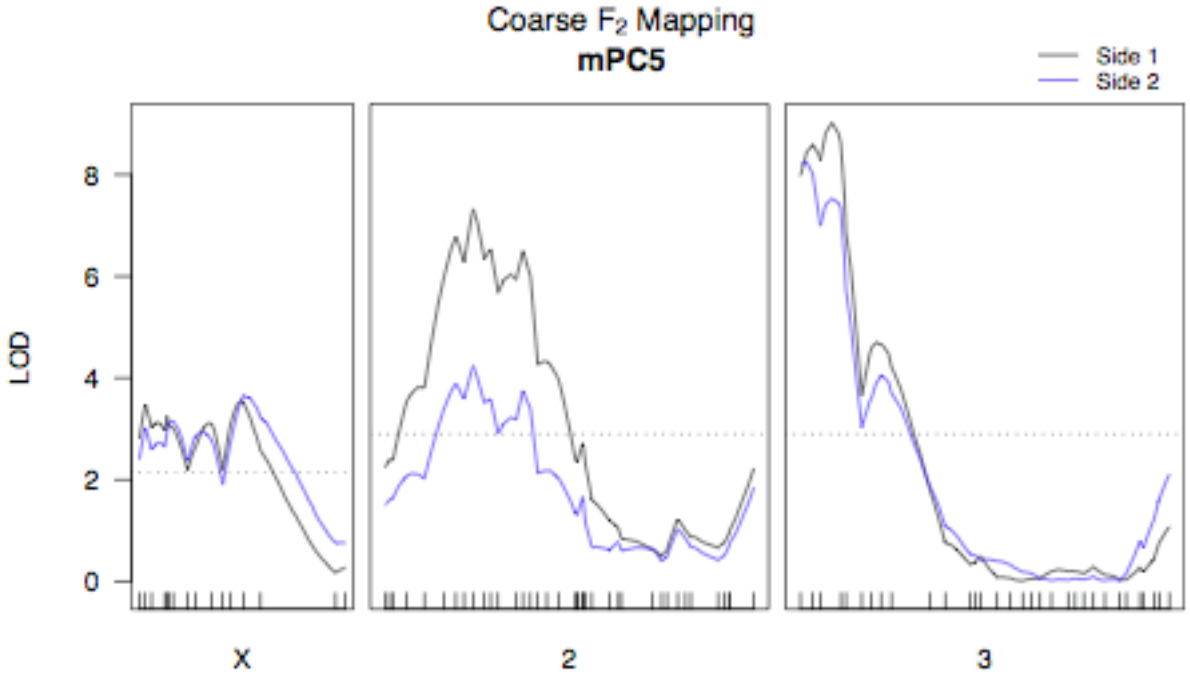

Coarse F2 Mapping
mPC4
Side 1
Side 2
LOD
15
10
5
0
X
2
3

Figure S3


Coarse F2 Mapping
mPC5
Side 1
Side 2
LOD
0
2
4
6
8
X
2
3

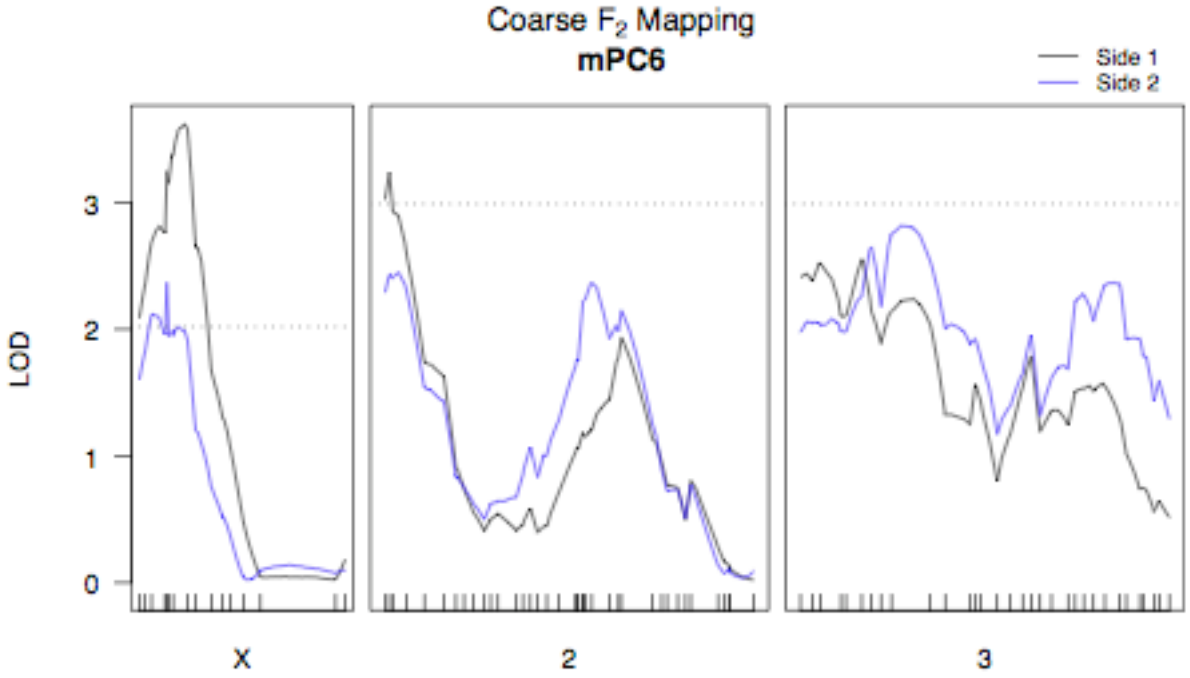
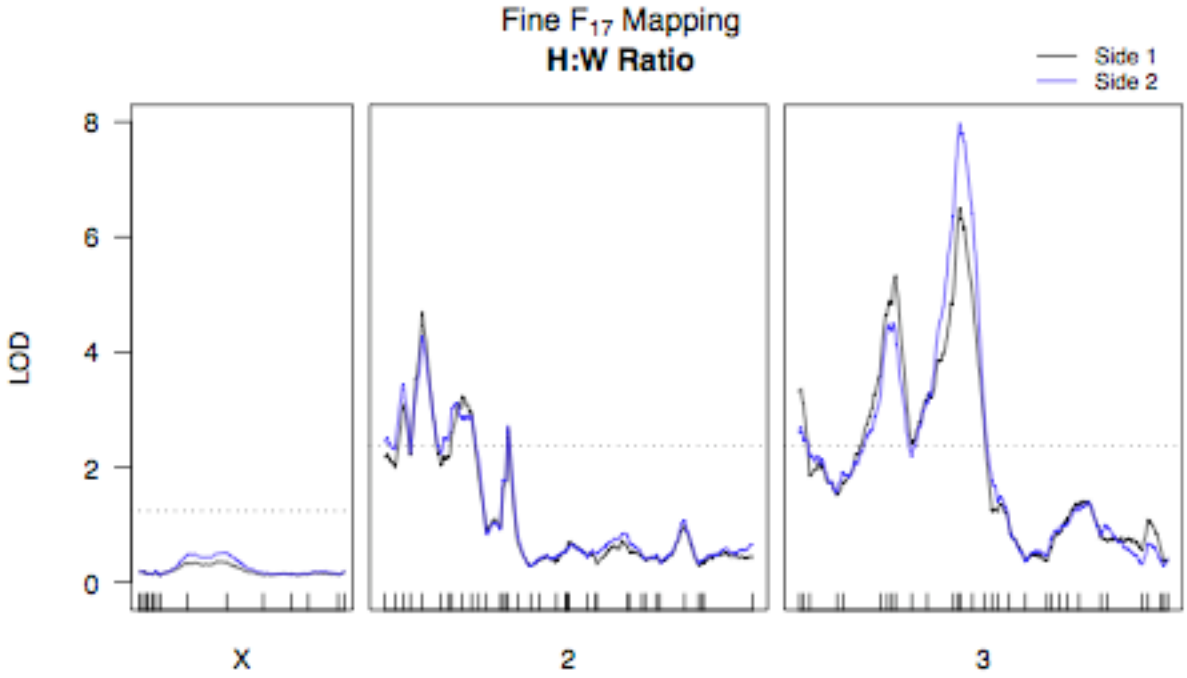

Coarse F2 Mapping
mPC6
Side 1
Side 2
LOD
0
1
2
3
X
2
3

Figure S3


Fine F_{17} Mapping
H:W Ratio
Side 1
Side 2
LOD
0
2
4
6
8
X
2
3

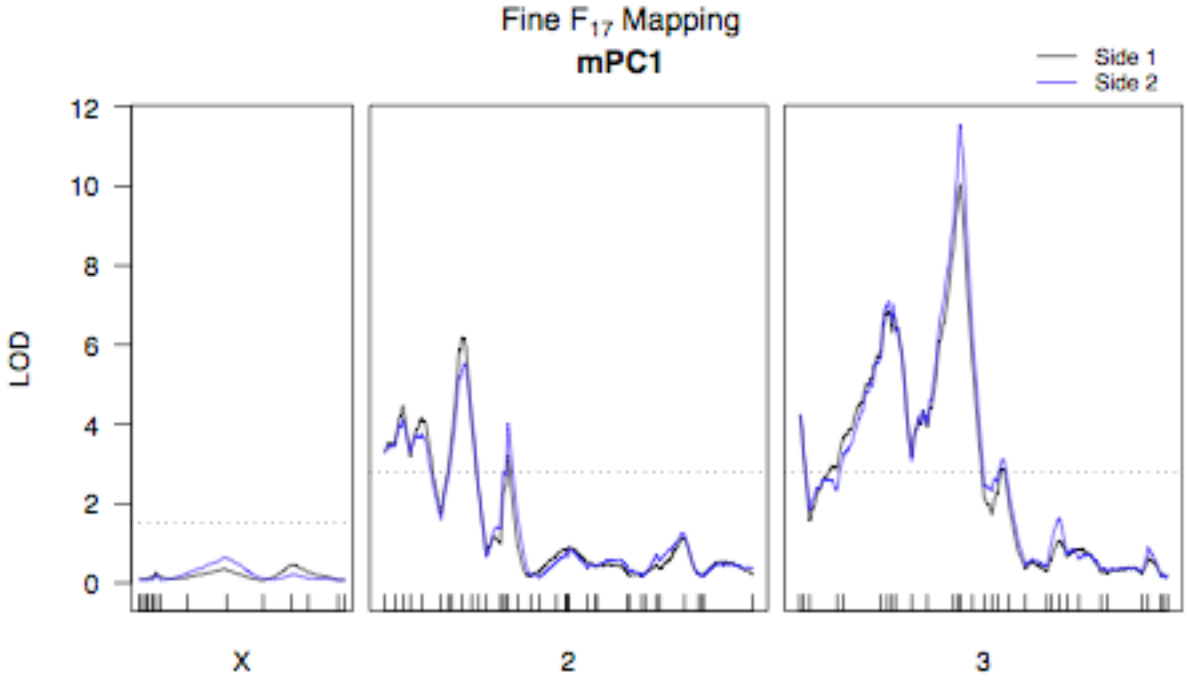

Fine F_{17} Mapping
mPC1
Side 1
Side 2
LOD
0
2
4
6
8
10
12
X
2
3

Figure S3 Coarse- and fine-mapping likelihood profiles for all traits. Each panel shows the results of interval mapping (IM) for a given phenotype, and the form of the plots is similar to those in Figure 4. Two curves are provided to demonstrate that similar results are generated no matter which of the two lobes is measured for a given fly. Note that the scale of the *y*-axis differs across plots.

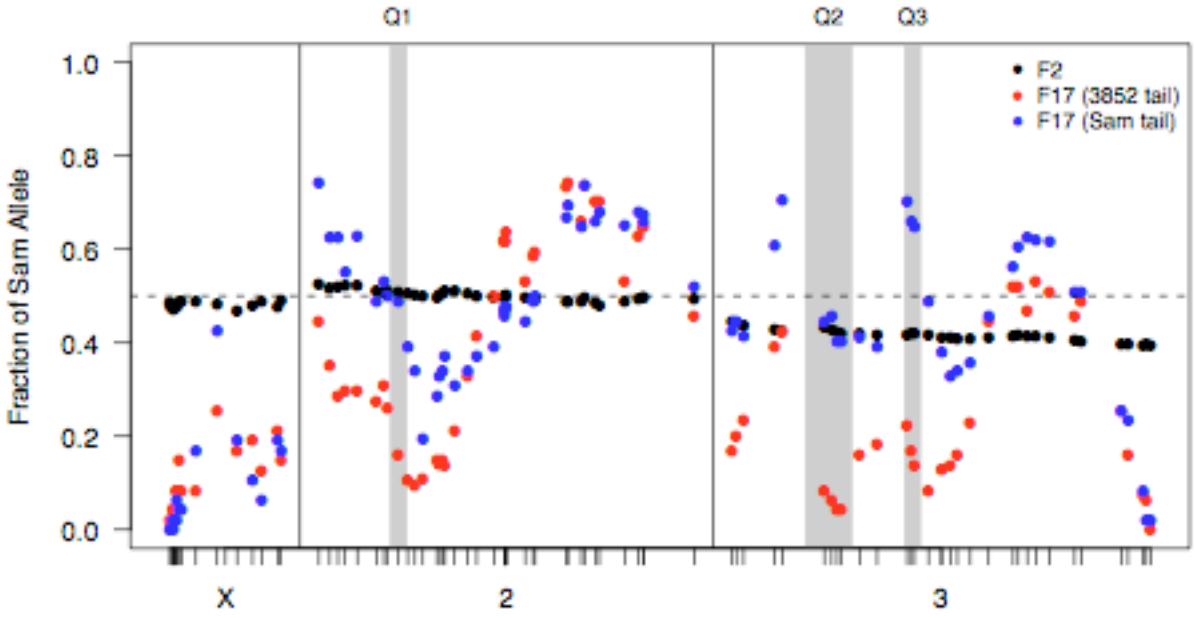


Figure S4 Frequency of the *Sam* allele at markers in the mapping panels. For every marker we directly count the fraction of alleles coming from the *Sam* parental line in each set of genotyped individuals - F_2 (black), b3852-like F_{17} tail (red), and *Sam*-like F_{17} tail (blue). The frequencies are plotted against the marker positions on the expanded F_{17} genetic map, and the three major QTL intervals are marked with gray boxes. In the F_2, *Sam* alleles are typically close to the expected frequency of 0.5, although frequencies are slightly lower for chromosome 3, potentially due to negative fitness consequences associated with the mutant ry^{506} allele. In the F_{17} both the X chromosome and the telomeric end of chromosome 3R show a dearth of *Sam* alleles in both tail samples, indicating mapping power in these regions is likely to be poor.

Chapter III

Quantitative trait loci for starvation resistance among inbred lines of the *Drosophila* Synthetic Population Resource

Abstract

Starvation is likely a major agent of natural selection in many animal species. While numerous studies have attempted to characterize the genetic contribution to starvation-resistance phenotypes in *Drosophila,* there has been relatively little success in identifying and validating genes that lead to quantitative variation in starvation resistance. To more thoroughly investigate the nature of the genetic variation controlling starvation resistance, we map QTL for starvation resistance using the *Drosophila* Synthetic Population Resource (King, *et al.* 2012). We starved 75,545 flies from 1725 RILs of the DSPR, facilitated by implementing high-throughput phenotyping strategies utilizing barcoded vials and semi-automated data collection. In the two populations of eight-way DSPR recombinant inbred lines (RILs) we identify at least 17 autosomal QTL. All QTL are specific to a single population of the DSPR and five are shared between males and females. The 2-LOD confidence interval around each QTL mapped to very small genetic and physical intervals (mean = 1.28cM and 0.91Mb) and encompass ~115 genes per QTL peak.

Introduction

Natural populations of animals commonly encounter periods of environmental stresses, such as desiccation, temperature, and starvation. The ability to endure short intervals of these and other common stresses form important life history traits. From a genetic perspective, the significance of stress resistance is clear, as several large families of genes (including heat-shock and cytochrome p450 genes) have evolved to deal with both specific and general environmental stresses. Though we possess some knowledge of generalized stress resistance across taxa, we have yet to fully characterize natural variation affecting stress resistance in any one species. Identifying specific loci that contribute to the genetic variation in stress resistance will help us gain an appreciation for the genetic basis of variation in life history traits and eventually allow us to test for evolutionary trade-offs that maintain genetic variation in nature.

Drosophila melanogaster is an outstanding model system to study the genetic basis of resistance to a multitude of stressors. There has been considerable success in characterizing quantitative genetic variation for heat and cold tolerance (Morgan and Mackay, 2006), bacterial load after infection (Lazzarro, Sackton, and Clark, 2006), and metabolism on several diets (Reed, et al. 2010). What is apparent from the body of work on stress resistance and life history traits in *Drosophila* is that the genetic architecture influencing each phenotype is highly polygenic, and there is likely to be some shared genetic architecture for stress resistance phenotypes and longevity in natural populations of flies.

When allowed access to moisture but not food, *D. melanogaster* isolates show substantial genetic variation in the time until death (starvation resistance phenotype) (Service and Rose, 1985, DaLage, *et al*. 1990, Hutchinson and Rose, 1991, and Hutchinson *et al.* 1991). Accordingly, starvation resistance has been a model quantitative trait for some time. Starvation resistance phenotypes show a robust response to artificial selection (Chippindale *et al*., 1996, Harshman and Schmid, 1999, Harshman *et al.* 1999a, and Sorensen, *et al*. 2007), and selection for starvation resistance has been positively correlated with increased lifespan (Hoffman and Parsons, 1993, though see Harshman, et al. 1999b) and resistance to environmental stressors including desiccation resistance (Rose, *et al*. 1992, Harshman, *et al.* 1999b), indicating that there may be generalized stress response mechanisms that function to maintain life during periods of environmental stress. Whilst being robust to environmental stressors is clearly important, there are potentially detrimental fitness consequences for animals that are extremely resistant to environmental stress. For example, selection for increased starvation resistance has been negatively correlated with fecundity (Rose, *et al*. 1992, Leroi, *et al*. 1994).

Several quantitative genetic mapping approaches have implicated both broad genomic regions and single genes in starvation resistance paradigms. Vieira, *et al*. 2000, utilized a set of 98 recombinant inbred lines (RIL) derived from cross between Oregon-R and 2b (Nuzhdin *et al*. 1997) to map QTL for lifespan under several environments, identifying five QTL for lifespan during starvation. Building on that work, Harbison, *et al*. 2004, utilized deficiency complementation testing (Pasyukova *et al*. 2000) to further refine the five QTL originally identified by Vieira, while also identifying a

set of *P*-element insertions that were significantly associated with starvation resistance. These studies implicated a number of candidate genes for starvation resistance that are involved in seemingly disparate biological processes including feeding behavior, cell fate specification, and cellular metabolism. The highly polygenic nature and unknown molecular underpinnings of stress resistance makes integrating potentially causative genes into an intelligible framework difficult, though not impossible. For example, Harbison et al. 2004, utilized observations that increased lipid levels are positively correlated with starvation resistance (Chippindale, *et al.* 1996, Harshman, *et al.* 1999a) to make a functional argument supporting the identification of *spalt major* (*salm*) as a gene involved in starvation resistance. *salm* had been previously implicated as a key gene in the development of oenocytes (Elstob, *et al*. 2001), specialized cells that play a major role in lipid metabolism (Gutierrez, *et al*. 2007) in *Drosophila*. Other studies have identified gene mutants that influence starvation resistance phenotypes as well, including genes involved in the insulin signaling pathway (Clancy, *et al*. 2001), autophagy (Juhasz, *et al*. 2007), and translation regulation (Tettweiler *et al*. 2005). As of yet, however, no quantitative genetic approach using naturally occurring mutations has implicated loci with such large-effects, indicating that the genetic architecture of starvation resistance in nature may be composed of alleles of somewhat subtle effects.

Recently, Mackay *et al.* 2012, in their introductory paper to the *Drosophila* Genetic Reference Population (DGRP), utilized an association mapping framework to identify SNPs that are associated with starvation resistance. Utilizing an arbitrary p-value cutoff of $P < 10^{-5}$ they find 203 SNPs associated with starvation resistance in a single population (flies of the DGRP were collected in Raleigh, North Carolina, USA).

Though the resolution of the association mapping approach is enviable, by mapping with relatively small number of lines, the DGRP study is unlikely to capture many alleles of modest effect (Long and Langley, 1999).

In this study we utilize a newly created quantitative genetic resource -- the *Drosophila* Synthetic Population Resource (DSPR) -- to build upon our knowledge of the quantitative genetic architecture of starvation resistance. A set of ~1700 recombinant inbred lines (RILs) founded from 15 inbred strains of worldwide distribution, the DSPR allows us to map quantitative trait loci (QTL) with extremely high resolution, estimate the frequency of QTL, and provides a route to define potentially causative alleles. Phenotyping the entire DSPR panel for starvation resistance we identify many QTL of small effect, while simultaneously obtaining an estimate of both their allelic effect and frequency.

Materials and Methods

D. melanogaster stocks

DSPR Founders:

The DSPR was initiated from fifteen highly inbred *P*-element free and *Wolbachia* free lines. Isolated from locations around the world, the founder lines exhibit a wide spectrum of phenotypic and genotypic variation present in *D. melanogaster*. Fourteen of the founder lines were purchased from the Bloomington or San Diego stock centers (see King, *et al.* 2012a), and one, *Samarkand; ry*^{506}, was a gift from T.F.C. Mackay

(Lyman *et al.* 1996). The genome of each DSPR founder line has been sequenced to facilitate the ultimate identification of causative genetic polymorphisms.

DSPR Flies:

For a full-description of the DSPR see King, *et al*. 2012a. Briefly, the DSPR is a community resource of ~1700 recombinant inbred lines (RIL) divided from a pair of 8-founder synthetic populations (pA and pB). Each founder population was used to initiate two subpopulations, (pA.1, pA.2 and pB.1, pB.2) which were allowed to reproduce *en masse* for 50 generations, and used to initiate inbred lines. Individual RILs were generated by ~25 generations of full-sib mating. RILs were genotyped using RAD (restriction-site associated DNA) markers (*e.g*. Baird, *et al.* 2008), resulting in a total of 10,275 biallelic SNPs across the three major chromosomes of the *Drosophila* genome. A hidden Markov model (King, et al. 2012b) was used to infer the underlying haplotypic structure of each RIL from the set of SNPs and the founder genome sequences, yielding a probability that a genomic interval is derived from each of the eight possible founders.

Experimental Treatments:

Throughout all experiments all flies, unless otherwise stated, were housed in a temperature and humidity-controlled room, under constant light, at ~23^{0}C and 40-60% relative humidity. All flies were reared in 25 x 95 mm polystyrene vials containing 10ml of cornmeal-molasses-yeast (CMY) medium. The starvation diet consisted of 10mL 1.5% agar in 25 x 95 mm polystyrene vials, to deprive the flies of virtually all caloric

content, without desiccating. Starvation media was less than 48 hours old at the beginning of each starvation treatment. Both CMY and starvation media were supplemented with 0.2% tegocept, 0.5%propionic acid, and 0.05% phosphoric acid, to prevent fungal and bacterial growth.

Experimental Workflow:

Duplicate vials of DSPR RILs were generated in blocks of 300 - 500 from the RIL stocks. To maintain a relatively constant and low larval density, DSPR flies were allowed to lay eggs for 24-48 hours and were discarded after initiating the experimental generation. Flies entering the starvation treatment were collected in single-sex groups of ten, 2-4 day old, non-virgin flies using CO_2. Collected flies were placed overnight in a 25 x 95 mm vial with CMY medium to recover from the gas treatment and feed. After 24-48 hours of recovery experimental flies were transferred to vials of starvation media. Starvation vials were placed in translucent Plexiglas trays, ensuring that, when shelved vertically, all vials were exposed to approximately equivalent amounts of light. Beginning 24 hours after the onset of starvation, the number of dead animals in each vial was recorded every 12 hours until all flies were dead. A total of 1725 RILs were assayed (mean = 2.22 vials/sex/RIL). Replicate vials were generated from the RIL stocks in independent experimental blocks.

Starvation Data Collection:

To facilitate high-throughput data collection from the DSPR, each RIL stock has a permanent, unique 5-digit barcode corresponding to the RIL number. We designed

custom Microsoft Excel macros that allowed us to use barcode scanners (model SC5, idautomation.com) to associate randomly-barcoded vials with DSPR RIL barcodes, effectively blinding the experimenter as to RIL identity for the duration of the experiment. The random barcode on each experimental vial was subsequently associated with two random 'starvation' barcodes (one for each sex of flies) at the time of starvation initiation, allowing us to efficiently collect and log data from >900 vials every 12 hours using a handheld barcode scanner (model HT630, Unitech).

QTL Mapping:

Mapping of QTL for starvation resistance follows the procedure outlined in King, et al, 2012a. Briefly, we regressed the median starvation death time for each RIL onto the additive probabilities that a RIL's genotype is derived from any of the eight possible founders at imputed marker loci spaced 10kb apart (highly similar results were obtained for mean starvation death time). Having observed that the starvation resistance phenotypic distributions of subpopulations pA.1 and pB.1 females differs significantly from that of subpopulations pA.2, and pB.2, respectively, we included subpopulation as a covariate in the regression both populations and sexes. We calculated LOD scores from the resulting ‘F’ statistic of the regression analysis (Broman and Sen, 2009), and further localized the peak of each QTL through standard interval mapping applied locally to the genomic region around each regression-based QTL peak (Lander and Botstein, 1989). We estimate the significance threshold for each sex and population using 1000 permutations of the phenotypic data (Churchill and Doerge, 1994) and define our QTL interval as the chromosomal region with LOD scores within 2 of each local peak.

Several QTL fell below our genomewide significance threshold upon localized interval mapping; for those we define both the QTL peak and 2-LOD confidence interval using data from the regression analysis.

QTL Phasing:

Our QTL phasing strategy relies on the assumption that, for any causative locus, two alleles reside among the eight founders of the population in which the QTL was identified. Accordingly, we estimate the phenotypic contribution from each founder haplotype at the peak of our QTL, and identify the shared allelic effects among founders (Macdonald and Long 2007, King *et al*. 2012a). When founder allelic effects at a QTL are not in two clearly defined groups, we enforce biallelism by only considering only the most phenotypically distinct founder lines, classifying the founders contributing 'intermediate' mean phenotypes as having unknown QTL phasing. Alleles that differ between 'high' and 'low' founder lines, define a series of putatively causative polymorphisms for each QTL (King, *et al*. 2012a).

Estimation of Heritability:

We estimated the broad sense heritability our starvation resistance assay by dividing the among RIL variance by the sum of the variance among RILs and the variance within RILs. To determine the heritability of RIL mean phenotypes, we use restricted maximum likelihood (REML) as in King, et al. 2012a.

Desiccation:

20 RILs from pA were selected from both of the high and low tails of the median female starvation resistance phenotypic distribution. Flies entering the desiccation treatment were collected identically to those exposed to starvation. To begin desiccation, each group of 10 flies was transferred into an empty vial sealed with a single layer of cheesecloth. Once all vials were populated, the vials were placed in a sealed 12"*h* x 12"*w* x 12"*d* Plexiglas box containing ~2.5lbs of Drierite desiccant. The number of dead flies in each vial was recorded every hour until all flies were deceased. Relative humidity (RH) dropped rapidly upon closure of the box. After 30 minutes RH was less than 10%, and was below 5% after 60 minutes.

***Drosophila* Activity Monitor Starvation:**

The high and low pA tail RILs used for desiccation were also assayed for activity and sleep of individual flies in the *Drosophila* Activity Monitor System (DAM) (www.trikinetics.com). Parental and experimental flies were reared in an incubator at 25^0C, 50%RH, and 12hr:12hr light:dark cycle. Experimental flies were collected in single sex groups of ~20, 2-3 day old male and female flies on CO_2 and allowed to recover overnight in vials containing CMY medium. 16 flies from each RIL were aspirated individually into DAM tubes containing a small plug of food (similar to CMY but with sucrose substituting for molasses) and sealed with a foam plug. After ~48 hours of acclimation to the DAM tubes, each fly was tipped into another DAM tube containing 1.5% agar starvation media. The activity level (number of infrared beam crossings) of each fly during starvation was recorded every minute and binned over 5-minute

intervals to observe bouts of activity or sleep. Any fly not crossing the IR beam in a 5-minute window was classified as sleeping (Shaw, *et al.*, 2000).

Results:

Variation for starvation resistance among the DSPR founder lines:

In order to characterize the phenotypic variation that exists among the 15 DSPR founder lines for starvation resistance, we undertook a pilot experiment (Fig. 1). Two to five replicate vials of 10 single-sex flies of each DSPR founder were starved (mean=4.83 ± 0.648 vials/founder/sex). DSPR founders show striking variation in mean starvation resistance, with line means ranging from 43.7hrs - 156.8hrs for males and 46.3hrs - 213.1hrs for females. In agreement with previous work on starvation, (Service and Rose, 1985, Harshman, *et al*., 1999a, Vieira *et al*., 2000) we find females to be longer-lived under starvation conditions than males (t-test: $p = 0.035$).

Female and male DSPR founder starvation means are highly correlated ($r^2 = 0.86$, $p = 3.316e\text{-}5$). Rank order of the founders between the sexes is correlated as well (Spearman's rho = 0.846, $p = 6.005e\text{-}05$). Despite the general congruency between male and female starvation, the lack of perfect rank-order correlation between the two sexes implies, as previously observed, (Vieira, *et al.* 2000, Mackay, *et al*. 2012) that the genetic contribution to starvation resistance is not identical between the sexes.

Starvation resistance among RILs of the DSPR:

To identify genomic regions contributing to starvation resistance among natural isolates of *D. melanogaster* we starved males and females of the DSPR. Figure 2 shows the extent of the phenotypic variation among 1725 RILs of the DSPR for mean starvation resistance. Mean starvation resistance for RILs of the DSPR ranges from 63.61hrs to 260.28hrs for females and from 67.52hrs - 215.37hrs for males. The RILs display slightly greater phenotypic variation than the founder lines, though the least resistant RILs are more resistant that the least resistant founders, possibly due to the purging of deleterious alleles during the population maintenance and inbreeding phases of RIL generation.

The distribution of female DSPR phenotypes is not significantly different than female founders (t-test: $p = 0.1293$), while males of the RILs are slightly more resistant than founder males (t-test: $p = 0.0031$). As in the founder variation study, females are more resistant to starvation than males (t-test: $p < 2.2e^{-}16$), and the correlation between male and female starvation within RILs is strong ($r^2 = 0.75$, $p < 2.2e^{-}16$). Rank order among mean starvation death times for males and females is relatively well preserved (Spearman's rho = 0.74, $p < 2.2e^{-}16$). Comparing starvation phenotypes of the RIL subpopulations (pA.1 vs. pA.2 and pB.1 vs. pB.2) we find that females of pA.1 are less resistant than those of pA.2 (t-test: $p = 1.463e^{-}05$) and females of pB.1 are less resistant than females of pB.2 (t-test: $p = 0.02834$). Mean starvation resistance phenotypes of males of pA.1 and pA.2 were not significantly different at the 5% level, as was the case for males of pB.1 and pB.2 as well.

Estimates of broad sense heritability (H^2) for mean starvation resistance are similar for both populations and sexes of the RILs (H^2 for males: pA = 0.519, pB = 0.550, for females: pA = 0.547, pB = 0.541). These values are in close agreement with the estimated heritability of starvation resistance among 168 inbred lines of the DGRP (H^2 = 0.581) using a relatively similar assay (Mackay, *et al*. 2012). We note, however that our estimate of H^2 is substantially higher than that of 0.20 from a set of 16 isofemale lines (Carrillo and Gibson, 2002) and much lower than a realized heritability estimate obtained through artificial selection h^2 = 1.141+/- 0.285 (Service and Rose, 1985).

To facilitate data collection in future experiments; we explored utilizing median starvation-induced death time for each RIL as our starvation resistance phenotype. Mean and median starvation-induced death times are highly correlated for both males (r^2 = 0.984, $p < 2.2e^{-}16$), and females (r^2 = 0.987, $p < 2.2e^{-}16$). Given this high correlation, we perform all QTL mapping analyses on the median starvation resistance for each RIL.

pA Female Tail desiccation resistance:

Reckoning that our measurement of starvation resistance could be correlated with other stress resistance traits, we decided to assay desiccation resistance among RILs of extreme starvation phenotypes. Selecting ~20 RILs from the most and least resistant tails of the pA female median starvation distribution, we discovered that highly starvation resistant RILs are more resistant to desiccation than RILs that succumb quickly to starvation stress (t-test: $p = 2.854e^{-}06$) (supp. fig. 1). This result suggests that

genetic variation in the DSPR for resistance to starvation stress also plays a significant role in resistance to desiccation. Indeed, artificial selection for increased resistance to starvation stress has been correlated with resistance to multiple environmental stressors, including desiccation, ethanol, and acetone (Harshman, 1999a), and gene expression analyses following selection regimes for increased longevity, desiccation, and starvation stress show substantial overlap in the levels of transcriptional regulation (Sorensen, et al. 2007).

pA Female Tail Starvation Activity:

Pursuing a more mechanistic understanding of starvation resistance, we assayed the activity level of the tails of the pA female median starvation in the *Drosophila* Activity Monitor System (DAM). We observed 16 female flies from each selected RIL in the DAM at 25^0C under a 12hr:12hr light:dark cycle, as opposed to 23^0C constant light of our vial starvation conditions. Despite the increased temperature, light cycle, and the flies being housed individually in DAM tubes, the pA tail flies remained phenotypically distinct (data not shown). In the DAM, the pA 'resistant' tail mean is 136.40hrs and the pA 'susceptible' tail mean is 78.34hrs (t-test: $p = 9.435e^{-}10$). Activity levels of female flies during starvation treatment suggest that increased activity among susceptible lines may contribute to their rapid decline (supp. fig. 2). This trend of increased activity among poorly resistant lines appears to be maintained throughout the lifetime of flies in the starvation treatment though we caution that, lacking a robust statistical test for activity during starvation, this conclusion may be somewhat tenuous.

QTL for Starvation Resistance

Utilizing 1646 RILs of the DSPR (pA = 801; pB = 845) to map QTL for starvation resistance we discover 17 autosomal QTL for starvation resistance among both populations and sexes (Fig. 3 and Table 1). No QTL were identified on the X-chromosome. Defining putative QTL location as the area of the genome encompassed by a 2-LOD reduction from each above-threshold peak, the average QTL interval is ~0.9Mb, a physical distance corresponding to 1.28cM on an F2 genetic map (Table 1). We consider any QTL with overlapping 2-LOD confidence intervals to be the same QTL. Using this rule we identify five QTL that are coincident in males and females of a single population, and one QTL shared among males of pA and pB.

QTL effect sizes are small, explaining between 4.15% and 14.28% (mean = 6.01%) of the H^2 of the mean RIL phenotype (Table 1). Though most of the identified QTL are of modest effect, together they explain a substantial portion of our heritability estimate for each population and sex. QTL discovered in pA males account for 48.29% of the estimated H^2, with QTL in pB males explaining 71.75% of the H^2. Similarly, pA female QTL explain 51.18%, and pB female QTL explain 65.94% of the estimated H^2.

Five of the 17 QTL we identified were detected in both males and females, indicating that male and female flies generally resist starvation through different genetic mechanisms. Indeed, sex specific loci influencing starvation resistance have been identified in both QTL mapping (Vieira et al. 2000) and in P-element insertion lines (Harbison, et al, 2004). The sizes and direction of QTL effects are typically consistent across sexes, (Table 1B).

Despite identifying 17 QTL within the DSPR, only one QTL was found to exist in both pA and pB -- pA2.1 and pB2.1 -- located near the telomere of chromosome 2L. This lack of coincident QTL is unexpected, as 7/17 QTL have more than one founder possessing the minor allele. Finding the at least two founders carrying the minor allele of a QTL in a single population indicates that, among the founders of that population, the minor allele is not rare, and likely exists in the complementary mapping panel. The lack of QTL that are coincident between the two mapping populations suggest that the genes contributing to starvation resistance among the two populations may not be identical. This suggests that an alternative physiological mechanism for starvation resistance may exist between the two populations.

Localized Association Mapping of Starvation Resistance

To gain a list of putatively causative polymorphisms, we investigated the number of 'in-phase' polymorphisms residing within our QTL intervals. There are a total of 2071 genes within the 2-LOD intervals of all QTL (numbers of genes under each peak range from 37 - 370). Of those, 1073 genes have at least 1 in-phase polymorphism residing within a transcript, greatly reducing the total number of genes that potentially harbor QTN for resistance to starvation.

Exploring the potential to gain functional information from combining the statistical power of QTL mapping with a high-resolution association mapping approach, we made use of the recently published association mapping result for resistance to starvation in the DGRP. Utilizing an arbitrary *p*-value cutoff of 10^{-5}, Mackay, *et al*. 2012

implicates a total of 203 SNPs (158 genes) for starvation resistance, only one of which (X:18,238,966) is formally significant under the Bonferroni adjustment for multiple testing. We searched for polymorphisms associated with starvation resistance according to the arbitrary *p*-value of 10^{-5} used in Mackay *et al*. 2012 that are also underneath a QTL. Only 11 of the 158 genes that are implicated by the DGRP study are coincident with a QTL identified in the present study (table 2).

We subsequently searched for genes within our QTL intervals that could be implicated in starvation resistance by Mackay, *et al*. 2012 in a localized association study. Performing this association study on the *p*-values of 305,675 SNPs identified in the DGRP that exist only within DSPR QTL intervals, we find only one SNP to be significantly associated with starvation resistance the DGRP, 2R:14,153,993. This locus is associated with female starvation resistance in the DGRP and lies within a QTL that is specific for female starvation resistance (pA2.4). Attempting to further reduce the statistical burden for multiple testing, we tested only SNPs within QTL intervals that were called in the DGRP and were found have either an in-phase SNP or reside within an INDEL among the DSPR founders. We then performed a Bonferroni statistical correction on the p-values of the remaining 16,573 DGRP SNPs. As before, this analysis again left us only one significant SNP (2R:14,153,993).

We reckoned that the substantial lack of overlap in the genetic architecture for starvation resistance between our dataset and that of the DGRP could be due to differences in our starvation phenotypes. The only ostensible differences in protocol are

the lower temperature at which the DSPR RILs were starved (23^0C) and the constant light environment in our assay. To explore the potential phenotypic consequences of these differences, we starved a sample of the most and least resistant lines of the pA female median starvation distribution in DAMS at 25^0C, and 12:12 light:dark cycle. The phenotypes of the pA tail RILs starved in the DAMs is highly correlated with RILs starved at 23^0C and constant light --among pA females r^2 = 0.870; p = 1.213e$^-$11 (male values are highly similar, data not shown). Thus it appears as though the lack of substantial overlap between genomic regions implicated in our study, and SNPs associated with starvation resistance in the DGRP is unlikely to stem from differences in experimental protocol.

Discussion

This study is the first QTL mapping study to dissect starvation resistance on a panel of flies derived from worldwide origins. Utilizing the highly recombinant nature of the DSPR genomes we map 17 QTL for starvation resistance to an average resolution of 1.28cM. Phasing DSPR founder alleles at QTL identifies 7 QTL to have common minor alleles (QTL with at least 2 founders possessing the minor allele), and 10 QTL in which only one founder can be positively identified as having the minor allele. Ascertaining the founder phasing at each QTL drastically reduced the number of candidate genes and polymorphisms for starvation resistance, facilitating the future study of these candidate loci.

The dearth of QTL that are coincident in pA and pB despite the large number of QTL identified is surprising. There are several scenarios that in our experimental

framework that could cause a lack of cross-population QTL. One plausible explanation is that the non-uniform distribution of founder genotypes at each locus (see fig. 2 King, et al. 2012a). As the two populations are each derived from eight founder lines (one founder is shared between the mapping populations), the expectation is that for any given locus 1/8th of the RILs genotype will be derived from each founder line. Our mapping panels often deviate from that expectation such that rather than phasing 15 founder haplotypes at each locus the mapping population averages ~12 founders per locus. This reduction of founder genome representation at each locus lowers the effective amount of population-level genetic variation in the RILs for each QTL. This reduced founder genome representation compromises our inability to phase a subset of founder genotypes at the QTL peak, due to some QTL completely missing founder genotypes or rarely occurring founders causing low-confidence estimates of allelic effects.

A second possible reason for the lack of cross-population QTL is that many causative alleles are rare in natural populations of *Drosophila*, and correspondingly are found in only one DSPR founder line. As our synthetic populations are were derived from *D. melanogaster* isolates from locations around the world, rather than from a single location, it is possible that our founder lines have a large number of 'private' alleles, i.e. alleles not shared by any other founder line.

A third scenario that could preclude discovery of cross-population QTL is that QTL identified in one population may be sensitive to the effects of genetic background. The ability of genetic background to profoundly effect quantitative trait variation has been widely reported, including in studies of both *Drosophila* wing shape (Dworkin, et al.

2009) and lifespan (Leips, and Mackay, 2000). Should QTN be shared between RILs of pA and pB, the vastly different genetic backgrounds existing between the two populations may alter or even eliminate direct genetic effects identified in the opposite population.

Another interesting observation is the striking lack of overlap between the QTL identified in this study and the SNPs implicated by Mackay, et al. 2012. This apparent lack of coincident loci may be caused by the modest effect size of QTL discovered in this study (mean effect size = 6.01%). Assuming the distribution of allelic effects in the DSPR is similar to that in the DGRP, the DGRP will struggle to identify loci contributing to starvation stress resistance. Indeed, an association mapping panel with a modest number of lines has low power to detect true associations for alleles contributing 5% to the additive genetic variance, (Long and Langley, 1999) and may also struggle to reject false positive associations.

Another plausible explanation for the lack of coincident loci between the DSPR and the DGRP is the possibility that starvation resistance alleles are truly rare in natural populations of *Drosophila.* This possibility is somewhat unlikely however, as the mean minor allele frequency for loci implicated in starvation resistance in the DGRP is 0.12. In the present study as well, we often find a 'rare' minor allele, with 10/17 QTL having only 1 founder line being identified as having the minor allele, though this observation is subject to the inability to phase several founder genotypes at each QTL, reducing our ability to precisely characterize the commonality of alleles among the DSPR founder lines.

A remaining possibility is that the genetic architecture of starvation resistance actually differs between the DSPR and DGRP. Indeed, the founder lines of the DSPR were harvested from around the world before *P*-elements swept through the *D. melanogaster* population and have been adapted to the laboratory for many years. In contrast, the DGRP consists of lines derived from a single location, and have been in laboratory culture for much less time. Thus it seems possible that there are substantial differences in genetic architecture of starvation resistance between the DGRP and DSPR.

Despite our high mapping resolution, we have not, as of yet, been able to localize starvation resistance QTL to a small set of candidate genes. Importantly, most of the QTL identified are of small effect sizes, signifying the need for the continued use of high power genetic mapping techniques for the analysis of life history traits in *Drosophila*.

Literature cited

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker E.U., Cresko, W.A., and Johnson, E.A. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* **3:** e3376.

Broman K.W. and Sen S. 2009. *A Guide to QTL Mapping with R/qtl*. Springer, New York.

Carrillo, R., and Gibson, G. 2002. Unusual genetic architecture of natural variation affecting drug resistance in *Drosophila melanogaster. Genet. Res., Camb*. **80**:205-213

Chippindale, A.K., Chu, T.J.F., and Rose, M.R. 1996. Complex trade-offs and the evolution of starvation resistance in *Drosophila melanogaster*. *Evolution* **50**:753-66.

Churchill, G. A., and Doerge, R.W. 1994. Empirical threshold values for quantitative trait mapping. *Genetics* **138:** 963–971

Clancy, D.J., Gems, D., Harshman, L.G., Oldham, S., Stocker, H., Hafen, E., Leevers, S.J., and Partridge, L. 2001. Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein. *Science* **292**: 104-106

Da Lage, J. L., Capy, P., and David, J.R. 1990 Starvation and desicca-
tion tolerance in *Drosophila melanogaster* : differences between European, North African, and Afrotropical populations. *Genetics Selection Evolution* **22**: 381–391.

Dworkin, I. Kennerly, E. Tack, D. Hutchinson, J. Brown, J. Mahaffey, J and Gibson, G. 2009. Genomic consequences of background effects on scalloped mutant expressivity in the wing of *Drosophila melanogaster*. *Genetics* **181**: 1065-1076.

Elstob, P.R., Brodu, V., and Gould, A.P. 2001. spalt-dependent switching between two cell fates that are induced by the *Drosophila* EGF receptor. *Development* **128**(5): 723--732

Gutierrez, E., Wiggins, D., Fielding, B., and Gould, A.P. 2007 Specialized hepatocyte-like cells regulate *Drosophila* lipid metabolism. *Nature* **445**:275-280

Harbison, S.T., Yamamoto, A.H., Fanara, J.J., Norga, K.K., and Mackay, T.F.C. 2004. Quantitative trait loci affecting starvation resistance in *Drosophila melanogaster* *Genetics* **166**: 1807–1823

Harshman L.G., and Schmid. J.L., 1998. Evolution of starvation resistance in *Drosophila melanogaster*: Aspects of metabolism and counter-impact selection. *Evolution* **52**(6): 1679-1685

Harshman, L.G., Hoffmann, A.A., and Clark, A.G., 1999a Selection for starvation resistance in *Drosophila melanogaster*: physiological correlates, enzyme activities and multiple stress responses. *Journal of Evolutionary Biology* **12**:370-379

Harshman, L.G., Moore, K.M., Sty, M.A., and Magwire, M.M., 1999b Stress resistance and longevity in selected lines of *Drosophila melanogaster. Neurobiology of Aging* **20**:521-529

Hoffmann, A.A. and Parsons, P.A. 1993. Selection for adult desiccation resistance in *Drosophila melanogaster*: Fitness components, larval resistance, and stress correlations. *Biol. J. Linn Soc*. **48**: 43–54.

Hutchinson, E. W., and Rose, M.R. 1991 Quantitative genetics of postponed aging in *Drosophila melanogaster*. I. Analysis of outbred populations. *Genetics* **127**: 719–727.

Hutchinson, E. W., A. J. Shaw and M. R. Rose, 1991 Quantitative genetics of postponed aging in *Drosophila melanogaster*. II. Analysis of selected lines. *Genetics* **127**: 729–737.

Juhasz, G., Erdi, B., Sass, M., and Neufeld, T.P. 2007 Atg7-dependent autophagy promotes neuronal health, stress tolerance, and longevity but is dispensable for metamorphosis in *Drosophila*. *Genes & Development* **21**: 3061-3066

King, E.G., Merkes, C.M., McNeil, C.L., Hoofer, S.R., Sen, S., Broman, K.W., Long, A.D., and Macdonald, S.J.. 2012a. Genetic dissection of a model complex trait using the *Drosophila* Synthetic Population Resource. *Genome Research* (Online before print, April 2012)

King, E.G., Macdonald, S.J., and Long, A.D. 2012b. Properties and power of the *Drosophila* Synthetic Population Resource for the routine dissection of complex traits. *Genetics* (Epub ahead of print Apr 13, 2012)

Lander, E. S., and Botstein, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121:** 185–199.

Lazzaro, B.P., T.B. Sackton and A.G. Clark. (2006) Genetic variation in *Drosophila melanogaster* resistance to infection: a comparison across bacteria. *Genetics* **174**:1539-1554

Leips, J., and Mackay, T.F.C. 2000. Quantitative trait loci for lifespan in *Drosophila melanogaster*: interactions with genetic background and larval density. *Genetics* **155**:1773-1788.

Leroi, A.M., Kim, S.B., and Rose, M.R. 1994. The evolution of phenotypic life-history trade-offs: an experimental study using *Drosophila melanogaster*. *American Naturalist* **144**:661-676.

Long, A.D., and Langley, C.H., 1999. The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Research* **9**:720-731.

Lyman, R.F., Lawrence, F., Nuzhdin, S.V. and Mackay, T.F.C. 1996. Effects of single *P*-element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics* **143:** 277–292.

Macdonald, S.J., Long, A.D. 2007. Joint estimates of quantitative trait locus effect and frequency using synthetic recombinant populations of *Drosophila melanogaster*. *Genetics* **176**: 1261–1281.

Mackay, T.F., Richards, S., Stone, E.A., Barbadilla, A., Ayroles, J.F., Zhu, D., Casillas, S., Han, Y., Magwire, M.M., Cridland, J.M., et al. 2012. The *Drosophila melanogaster* genetic reference panel. *Nature* **482:** 173–178

Morgan, T.J., and Mackay, T.F.C. 2006. Quantitative trait loci for thermotolerance phenotypes in *Drosophila melanogaster.* *Heredity* **96**: 232–242

Nuzhdin, S.V., Pasyukova, E.G., Dilda, C.L., Zeng, Z.-B., and Mackay, T.F.C. 1997 Sex-specific quantitative trait loci affecting longevity. *Proc. Natl. Acad. Sci. USA* **94**: 9734-9739

Pasyukova, E.G., C. Vieira, and T.F.C. Mackay, 2000 Deficiency mapping of quantitative trait loci affecting longevity in *Drosophila melanogaster.* Genetics **156**: 1129-1146

Reed, L.K., Williams, S., Springston, M., Brown, J., Freeman, K., DesRoches, C.E., Sokolowski, M.B., and Gibson, G. 2010 Genotype-by-diet interactions drive metabolic phenotype variation in *Drosophila melanogaster. Genetics* **185**: 1009-1019

Rose, M.R., L.N. Vu, S.U. Park, & J.L. Graves. 1992. Selection for stress resistance increases longevity in *Drosophila melanogaster*. *Experimental Gerontology* **27**: 241-250

Service, P.M., and Rose, M.R., 1985. Genetic covariation among life-hisotry components: The effect of novel environments. *Evolution*, **39**(4): 943-945

Shaw, P.J., Cirelli, C., Greenspan, R.J., and Tononi, G. 2000. Correlates of sleeping and waking in *Drosophila melanogaster. Science* **287**: 1834-1837

Sorensen, J. G., Nielsen, M. M., and Loeschke, V., 2007 Gene expression profile analysis of *Drosophila melanogaster* selected for resistance to environmental stressors. *Journal of Evolutionary Biology* **20**: 1624-1636

Tettweiler, G., Miron, M., Jenkins, M. Sonenberg, N., and Lasko, P.F. 2005. Starvation and oxidative stress resistance in *Drosophila* are mediated through the eIF4E-binding protein, d4E-BP. *Genes & Development* **19**: 1840-1843

Vieira, C., Pasyukova, E.G., Zeng, Z.B., Hackett, J.B., Lyman, R.F., and Mackay, T.F.C. 2000 Genotype-environment interaction for quantitative trait loci affecting life span in *Drosophila melanogaster. Genetics* **154**:213-227

Figures and Tables

Table 1 (next page) Starvation resistance QTL. Panel A: QTL found in only one sex. Panel B: QTL identified in both males and females.

[a] QTL are numbered using their population, chromosome number, and order of peak position.

[b] The physical position (bp) of each QTL peak for flybase release **???**

[c] The genetic distance along a chromosome spanning a 2-LOD reduction in peak intensity

[d] The portion of H^2 of RIL mean phenotype for each sex/RIL population explained by the QTL

Table 1

	QTL Name[a]	Pop	Sex	Chr	Ppos[b]	Peak LOD	Distance (Mb)[c]	Distance (cM)[c]	QTL Effect[d]	No. of Genes
A)										
	pA2.1	A	m	2L	420000	7.22	0.29	0.635	4.61	44
	pA3.3	A	m	3R	15840000	7.25	0.31	0.842	4.63	47
	pA3.4	A	m	3R	21250000	7.75	0.60	2.144	4.94	112
	pB2.1	B	m	2L	270000	7.25	0.99	1.815	4.10	141
	pB2.3	B	m	2L	10470000	7.18	0.37	1.562	4.15	103
	pB2.6	B	m	2R	2080000	18.75	4.36	1.702	10.51	370
	pA2.3	A	f	2R	13380000	7.57	0.90	3.170	4.63	179
	pA2.4	A	f	2R	14710000	7.22	0.58	1.648	4.42	103
	pA3.1	A	f	3L	3650000	10.49	0.52	1.998	6.37	41
	pB2.2	B	f	2L	7110000	7.76	0.35	1.895	4.58	28
	pB2.5	B	f	2L	17450000	12.11	2.49	1.359	7.06	243
	pB3.3	B	f	3R	25940000	7.50	0.49	0.974	4.43	104
B)										
	pA2.2	A	m	2R	5080000	9.55	0.61	1.220	6.06	92
	pA2.2	A	f	2R	5240000	9.57	1.37	2.698	5.82	220
	pA3.2	A	m	3R	2300000	7.56	1.61	0.444	4.82	223
	pA3.2	A	f	3R	1180000	11.12	1.48	0.393	6.74	208
	pB2.4	B	m	2L	12200000	7.30	0.40	0.818	4.22	37
	pB2.4	B	f	2L	12320000	7.24	0.54	1.077	4.28	40
	pB3.1	B	m	3L	19110000	26.00	0.48	0.271	14.28	42
	pB3.1	B	f	3L	19060000	18.29	0.51	0.288	10.49	43
	pB3.2	B	m	3R	9160000	11.01	0.35	0.538	6.30	44
	pB3.2	B	f	3R	9160000	8.22	0.50	0.763	4.85	64
						Mean:	**0.91**	**1.284**	**6.01**	**115**

Table 2. Genes under QTL that are implicated in starvation resistance by the arbitrary *p*-value cutoff in the association study of Mackay *et al.* 2012

[a] Polymorphisms called in the DSPR that: 1) Match the pattern of founder allelic phasing and 2) Reside within the transcript of the implicated gene

[b] SNPs are considered significant if their p-value is less than the Bonferroni-adjusted 0.05 significance level for localized association mapping (see text)

[c] The number of founder means that are rare over the total number of founder means phased at a given QTL peak

DGRP SNP	DGRP p-value	QTL	Implicated Gene	In-Phase Polymorphism[a]	Significant Association[b]	Minor Allele Founder Freq. [c]
2L:841909	1.75e-06	pB2.1	*drongo*	Yes	No	1/3
2R:1743883	1.17e-06	pB2.5	*dpr12*	No	No	1/5
2R:6231292	6.92e-06	pA2.2	CG42732	No	No	1/5
2R:6231357	2.31e-06	pA2.2	CG42732	No	No	1/5
2R:13072510	1.16e-06	pA2.3	CG10950	Yes	No	1/3
			CG10953	Yes	No	
2R:14153993	7.84e-08	pA2.4	*Ote*	Yes	Yes	1/5
3L:3829645	8.38e-06	pA3.1	Awh	Yes	No	1/3
			enc	Yes	No	
			Rdh	Yes	No	
3L:19220699	3.27e-07	pB3.1	fz2	Yes	No	1/5
3R:25727333	3.12e-06	pB3.3	CG31038	Yes	No	3/6

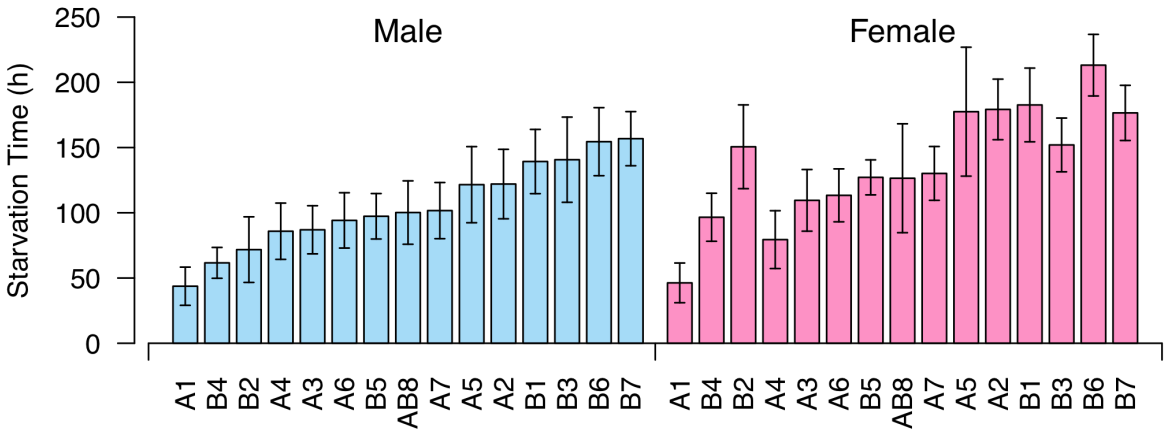


Figure 1. Mean (±SD) starvation resistance phenotypes of the 15 DSPR founder lines. Replicate vials of 10 male or female flies from each of the DSPR founder lines were starved on 1.5% agar until all flies were dead. Founder lines are ordered by the male mean starvation resistance phenotype. Note that rank order, although correlated, is not strictly preserved between sexes. For a complete description of the DSPR founder lines see King, et al. 2012a.

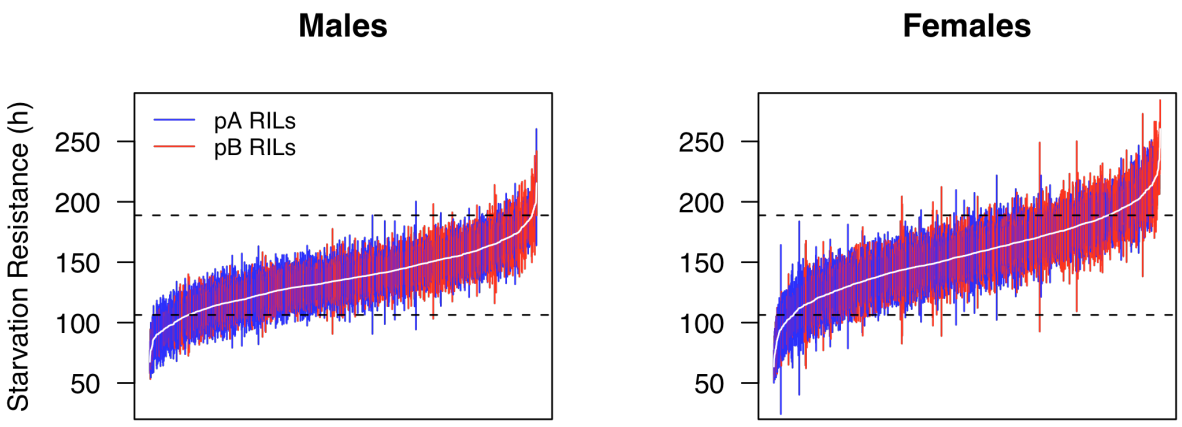


Figure 2. Distribution of starvation resistance phenotypes among 1725 RILs of the DSPR. The mean (white line) ± standard deviation (blue and red whiskers) starvation resistance phenotypes for both male and female flies are shown. Both male and female individuals from pB RILs are more resistant to starvation than are pA RILs (t-test: $p < 2.2e^{-}16$ for each sex). Dashed lines indicate the 5^{th} and 95^{th} percentiles of the combined male and female phenotypic distribution, and reveal that female *D. melanogaster* are more robust to starvation conditions than are males (see text).

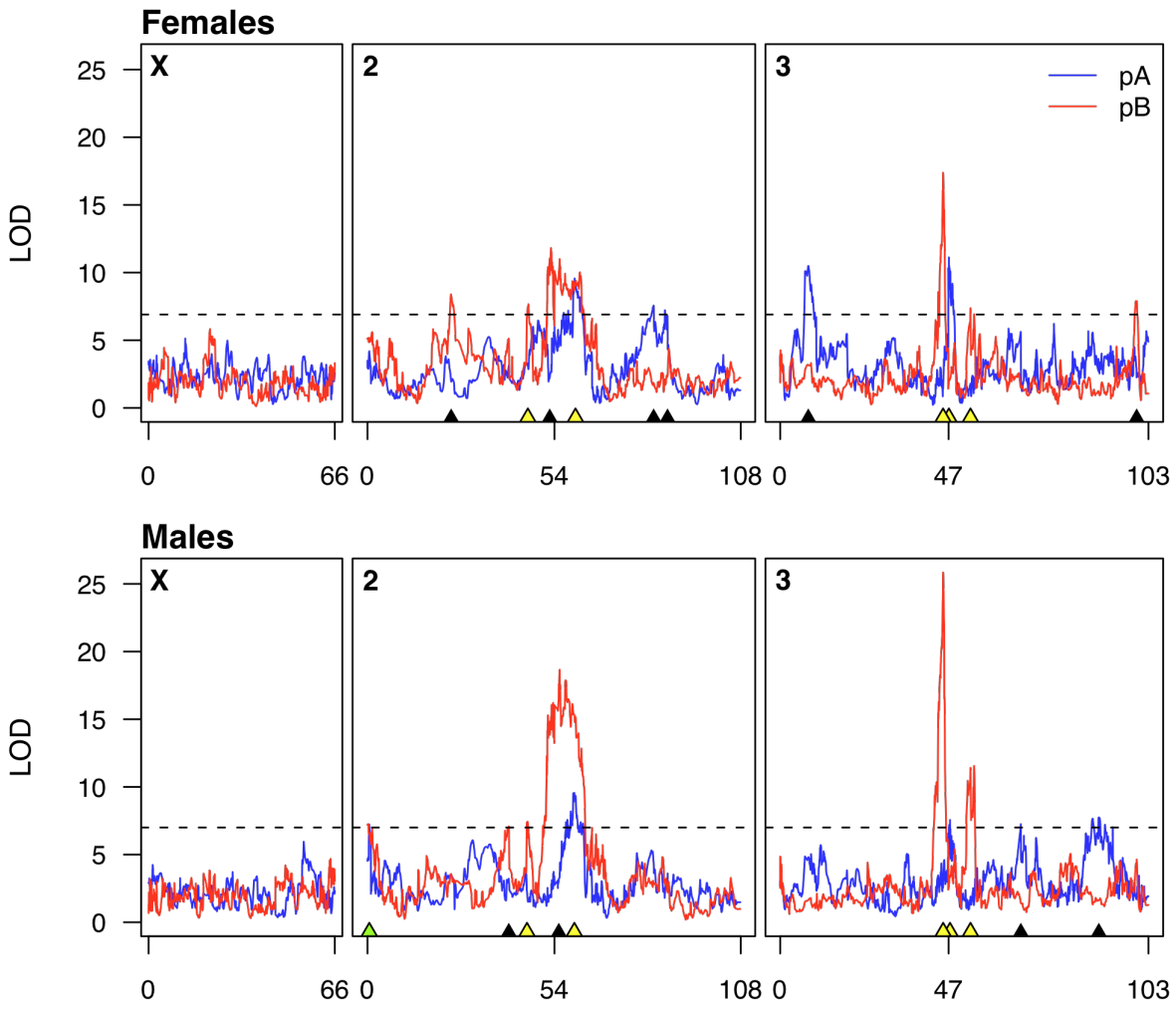


Figure 3. QTL for median starvation resistance among the DSPR. Chromosomes are indicated in the top left of each box. Blue (red) curves represent the likelihood that RILs of pA (pB) harbor QTL for starvation resistance. Genetic distance relative to the standard F_2 recombination map is shown on the *x*-axis for each chromosome. Genomewide significance levels were determined by 1000 permutations of the phenotypic data and are as follows: females pA = 6.9LOD, pB = 6.9LOD, and males pA = 7.0LOD, pB = 6.6LOD. Dashed lines indicate the more stringent LOD threshold for each sex. Triangles along the *x*-axis indicate each of the 21 QTL discovered during this experiment. Black triangles represent QTL found only in one sex, while yellow triangles represent QTL that were identified in both males and females (see Methods). No QTL were discovered in both pA and pB.

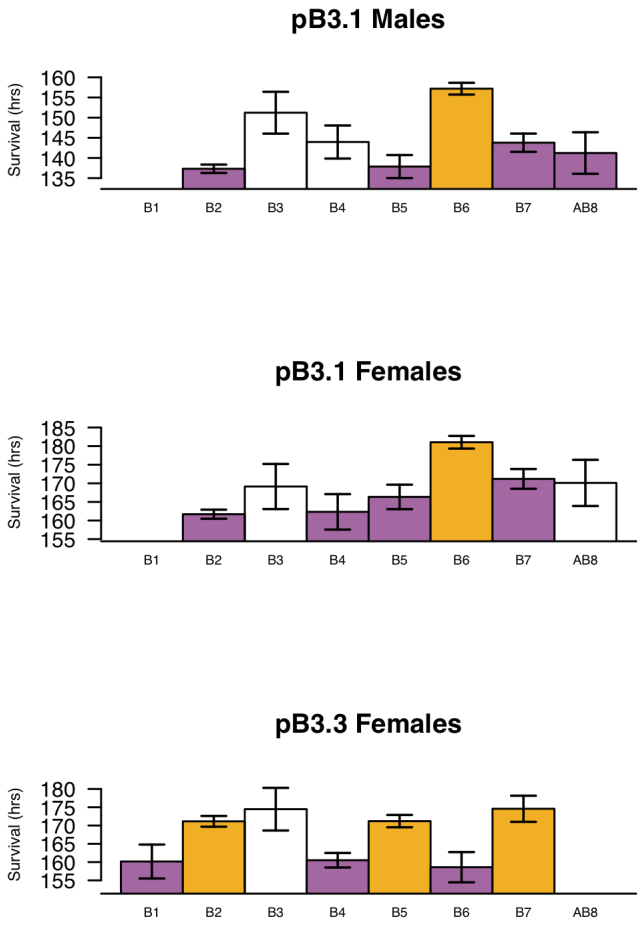
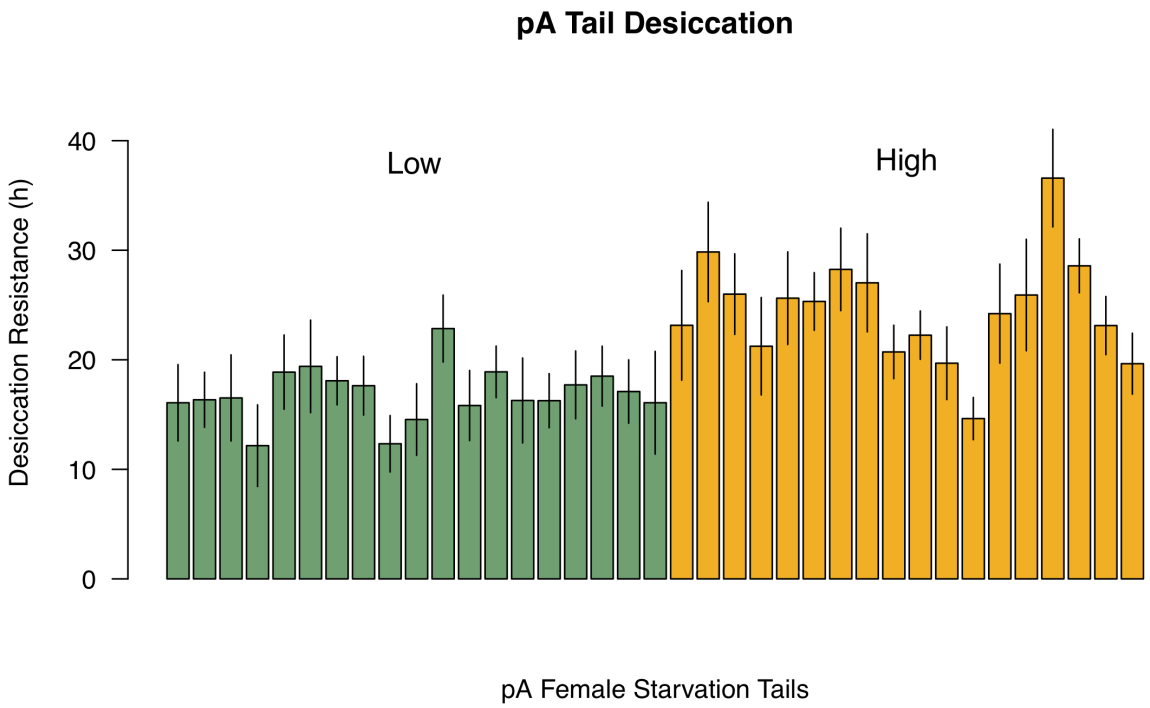


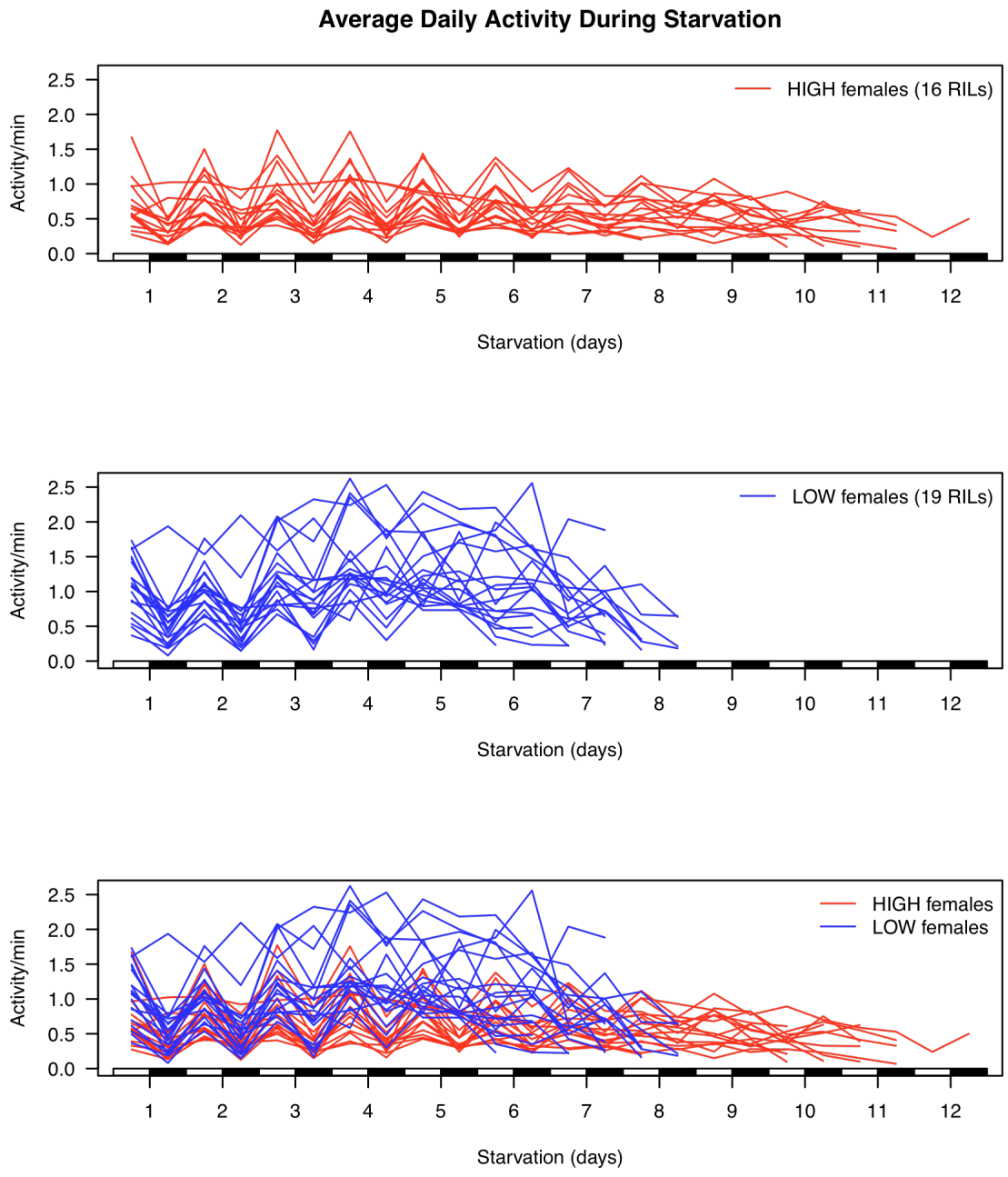
Figure 4. Founder mean contribution at selected QTL. QTL phasing for each of the founder genotypes is indicated. In all cases, founder line means in orange indicate that founder has the resistant allele, while founders represented in purple have the susceptible allele. White bars indicate the mean contribution of that founder is not phased (note the relatively large standard error around the founder means). Founder lines with no bar indicate that, at the physical location of the QTL peak, fewer than 5 RILs are derived from that founder line. QTL pB3.1 is found in both male and female RILs (panel A), with the resistant allele being unique among phased founder genotypes, and is presumably rare among D. melanogaster isolates. Panel B shows QTL pB3.3, identified only in females, with the resistant and susceptible alleles equally common among the phased founder genotypes.



Supplementary Figure 1. Mean (±SD) desiccation induced death time of female flies derived from either the poorly resistant (green bars) or strongly resistant (gold bars) tails of the pA starvation resistance distribution. Replicate vials of 10 female flies from each RIL were desiccated in empty vials until all flies were dead (see methods). RILs that are strongly resistant to starvation stress are longer lived under desiccation stress than are RILs that succumb quickly to starvation (t-test: p = 2.854e^{-}06).

Supplementary Figure 2 (next page). Activity profiles of starving female pA RILs in the DAM. RILs of the 'resistant' tail of the pA female starvation distribution are shown by red lines with RILs of the 'susceptible' tail shown in blue. RIL activity is estimated as the average activity level (number of beam crossings) per minute of all living flies during each 12hr light period (*x*-axis white bars) or 12hr dark period (*x*-axis black bars). Note that RILs of the susceptible tail die substantially faster than do the flies of the resistant tail. Also note the relatively high activity level of the susceptible flies as compared to the starvation resistant flies.

Supplementary Figure 2



Chapter IV

Mapping quantitative trait loci for starvation resistance in three outbred genetic backgrounds reveals a highly complex genetic architecture for this model life history trait.

Abstract

Although the *Drosophila* Synthetic Population Resource (DSPR) provides a number of advantages over prior linkage mapping and common association mapping approaches, there are two major concerns with relying on genetic mapping within inbred lines. First, inbred, homozygous genotypes are clearly not representative of naturally-derived flies. Second, mapping directly in inbred lines allows for the identification of QTL in only a small subset of genetic backgrounds, rather than the innumerable permutations of genetic backgrounds obtainable in nature. In light of these challenges, we mapped QTL for starvation resistance utilizing the DSPR in three outbred genetic environments. Mapping QTL in a series of round robin crosses between lines of the DSPR and backcrosses between DSPR lines and two isogenic strains, we discover a highly complex genetic architecture for starvation resistance, identify many cross-specific QTL, all of which have small effects. Notably, each QTL originally mapped in inbred lines of the DSPR was replicated in at least one cross, indicating that QTL mapped among inbred lines are specific to starvation resistance likely not due to inbreeding depression.

Introduction

The quantitative nature of the genetic variation underlying life-history traits has complicated the study of evolutionary and behavioral genetics. Since virtually all fitness traits, including viability, fecundity, and longevity are quantitative in nature, building an understanding of the genetic control of these traits is essential for understanding evolution of natural populations. Of particular interest is the development of strategies to understand the genetic architecture of natural populations, or of semi-natural laboratory populations that mimic the general properties of individuals found in nature. Do these naturally derived alleles act in an additive or dominant manner? Are they protein-coding or regulatory mutations? Only by gaining an estimate of these parameters in naturally derived genetic backgrounds will we begin to identify the genetic variability upon which selection acts.

Drosophila melanogaster is an outstanding model system to address these intricate genetic questions. Indeed, *Drosophila* has been widely used as a model system to characterize quantitative genetic variation for a number of life history traits. For example, QTL for lifespan have been identified in a number of experimental frameworks (Forbes, et al. 2004, Nuzhdin, et al. 2005, and Wilson, et al. 2006). In each case at least 10 QTL were identified, revealing the highly polygenic nature of the genetic control of lifespan. Other life history traits such as thermotolerance (Morgan and Mackay, 2006) and age-specific fecundity (Leips, et al. 2006) have also been examined in *D. melanogaster*, although the specific alleles influencing these phenotypes remain mostly unknown.

There are several reasons for our lack of a thorough characterization of life history traits in *Drosophila.* First, seminal studies broadly characterizing natural history traits (e.g. Rose and Charlesworth, 1981a), or examining the effects of artificial selection on life history phenotypes (Rose and Charlesworth 1981b, Harshman and Schmid 1998) were not designed to identify the precise alleles contributing to these traits. Second, studies purporting to identify natural allelic variation for life history traits can be statistically underpowered (e.g. Mackay, et al 2012), and therefore are unable to detect true, repeatable associations at alleles of small effects. Third, studies that have identified genomic intervals or specific alleles that influence life-history traits have removed these causative alleles from a natural genetic background -- generally through the use of inbred lines for genetic mapping (Vieira, et al. 2000, Mackay, et al. 2012, McNeil, et al. 2012a), or through deficiency mapping, in which natural alleles are assayed in *trans* with laboratory-derived chromosomal deletions (Pasyukova, et al. 2000, Harbison, et al. 2004).

To build on our knowledge of quantitative genetic variation for life-history traits in *Drosophila*, we have chosen to study starvation resistance. Like other life-history traits, starvation resistance is highly polygenic (Vieira, et al. 2000, Harbison, et al. 2004, McNeil, et al. 2012), with most identified QTL contributing modest effects. Interestingly, QTL for starvation resistance are often sex specific (Vieira, et al. 2000, Forbes, et al. 2006, McNeil, et al. 2012) and are often specific to a single genetic mapping population. While the body of knowledge existing for starvation resistance and other life history traits in *Drosophila* is relatively strong, we still have limited knowledge regarding the applicability of our estimates of genetic architecture to natural populations of animals.

Faced with these concerns, we sought to develop a genetic mapping strategy that would allow us to precisely map QTL for starvation resistance in a reproducible outbred population, while obtaining a wide sample of natural genetic variation, mimicing that found in nature (i.e. low levels of inbreeding and high heterozygosity). To accomplish these goals we crossed flies from population A (pA) of the *Drosophila* Synthetic Population Resource (DSPR) recombinant inbred lines (RIL) to other pA RILs, as well as to two independent, isogenic strains. Mapping QTL for starvation resistance, we compare the results from these crosses to the mapping results from the inbred pA RILs (McNeil, et al. 2012). Bolstering the mapping result of the inbred pA RILs, we replicate each QTL initially discovered and additionally identify many novel, cross-specific QTL for starvation resistance. Our study highlights the complex genetic architecture underlying life-history traits in outbred populations of *Drosophila,* and underscores the need for genetic mapping experiments to be performed in outbred animals.

Materials and Methods

Population 'A' DSPR Founders and RILs:

The founders of pA of the DSPR are eight *P*-element and *Wolbachia* free, highly inbred lines (see Supplementary Table 1 of King et al 2012a). At least one founder was isolated on each continent save Antarctica, and the founders display a broad range of phenotypic and genotypic variation. Each DSPR founder genome has been resequenced to facilitate the identification of putative causal alleles after QTL mapping.

pA is half of the DSPR, a community resource of ~1700 recombinant inbred lines (RIL) derived from a pair of eight founder synthetic populations. The eight 'A' founders were used to initiate a pair of replicate synthetic subpopulations -- pA.1 and pA.2 -- each of which were allowed to reproduce *en masse* for 50 generations. RILs were generated by ~25 generations of full-sib mating and were genotyped using RAD (restriction-site associated DNA) markers (*e.g*. Baird, *et al.* 2008). A hidden Markov model (King, et al. 2012b) utilizes the set of SNPs identified in each RIL and the founder sequences, to infer the underlying founder haplotypic structure of each RIL genome. This process yields a probabilistic estimate that a given genetic locus is derived from each of the eight possible founders.

pA x pA Round Robin Crosses

A round robin crossing design was implemented to generate F1 animals that are derived from two sequential RILs of pA. These crosses allow us to estimate the heterozygous effect of each founder allele in a largely outbred genetic environment. Additionally, this design nearly eliminates the possibility of mapping QTL due to inbreeding depression. Assuming that the founder lines are equally frequent at any given genetic position among RILs, the probability of generation of a homozygote at any genetic position is 0.125.

Recurrent Parents

We chose to utilize two well-studied inbred lines of *D. melanogaster* that carry recessive markers to use as recurrent parents for our backcross design experiments.

Oregon R, (Lindsley and Zimm 1992) possessing an active *Wolbachia* infection, was a kind gift by Ted Morgan (Kansas State University). We also utilized the *D. melanogaster* sequenced strain (*iso1*) (Bloomington stock number 2057; Adams, et al. 2000), which we cleared of a *Wolbachia* infection by rearing the flies on medium containing tetracycline. Stocks of (*OreR* and *iso1*) were kept at large census sizes in 0.5 gallon milk bottles and pint-bottles, and were distributed into 25 x 95 mm polystyrene vials for each block of the experiment to facilitate virgin collection of the parental generation.

Experimental Treatments:

Starvation treatments were carried out identically to those of McNeil, et al. 2012. All experimental replicates were kept under constant light at ~23^0C and 40-60% relative humidity. Flies of the parental and experimental generations were reared in 25 x 95 mm polystyrene vials containing 10ml of cornmeal-molasses-yeast (CMY) medium. The starvation treatment consisted of freshly prepared (less than 48 hours old) 25 x 95 mm polystyrene vials containing 10mL of 1.5% agar supplemented with 0.2% tegocept, 0.5%propionic acid, and 0.05% phosphoric acid. This treatment effectively deprives the flies of nutrition without desiccation stress.

Experimental Workflow:

For a schematic of our workflow see fig. 1. We duplicated vials of DSPR RILs in blocks of 60 - 200 from the RIL stocks and concurrently generated 100 - 150 vials of *OreR* and *iso1*. Flies were allowed to lay eggs for 24-48 hours to maintain a relatively

constant, low larval density across experimental vials and were discarded after initiating the parental generation. Virgin females from the recurrent parents and each RIL were collected on CO_2 into groups of 5 to 10 flies, and were held in vials for at least 24 hours to recover from anesthesia. Five replicate vials of five male flies were collected from each RIL and were subsequently crossed to virgin females of both recurrent parents and a different RIL. Parental flies were allowed to mate and oviposit for ~48 hours before being discarded.

In the next generation, mated, 2-4 day old, experimental F_1 female flies were collected in groups of ten using CO_2. These flies were placed overnight in a 25 x 95 mm vial with CMY medium to recover from the gas treatment and feed. After ~24 hours of recovery experimental flies were transferred into the starvation treatment. Experimental vials were placed in translucent Plexiglas trays to ensure all vials were exposed to approximately equivalent amounts of light. After 24 hours of starvation, the number of dead animals in each vial was recorded every 12 hours until at least 6 of the flies were dead. A total of 849 crosses between pA RILs, 849 crosses of pA RILs to *OreR,* and 779 crosses of RILs to *iso1* were phenotyped for starvation resistance.

For clarity, from this point on the four genetic backgrounds will be named as follows: Inbred RILs (Inbred), OreR x RIL cross (OreRx), Iso1 x RIL cross (Iso1x), and RIL x RIL round robin crossing (RIX).

QTL Mapping:

QTL mapping among crosses in this experiment is essentially unchanged from that of King, et al. 2012a. To map QTL, we assume that all QTL effects are additive. For the backcross design experiments we regress the starvation phenotype on the eight

additive genotypic probabilities (King, et al. 2012a). With the round robin crosses (RIX design) (Tsaih, et al. 2005), we average the maternal and paternal genotype probabilities, and regress the starvation phenotype onto the genotype of each RIX cross. In both cases, we convert the resulting F-statistic to a LOD score (Broman and Sen 2009), and determine a significance threshold through 1000 permutations of the phenotypic data (Churchill and Doerge 1994). Though several previous mapping experiments (including inbred RIL starvation resistance) have shown subtle but significant differences in the mean phenotype between pA.1 and pA.2 (likely due to underlying population structure) we do not include subpopulation in our regression model, which may result in slightly inflated test statistics.

QTL Phasing:

Our QTL phasing strategy relies on the assumption that any causative locus is biallelic among the eight founder lines of pA. At the peaks of QTL in each cross we estimate the mean starvation phenotype for each founder genotype. We group (or phase) founder haplotypes through their similar mean allelic effects on the mean starvation resistance phenotype (see King *et al*. 2012a). For the purposes of identifying in-phase polymorphisms for a given QTL, if there are more than two founder allelic effect groups, we enforce biallelism by considering only the most phenotypically distinct founder haplotypes. Alleles that differ between 'high' and 'low' founder lines -- with the remaining founder lines effectively considered to be uninformative -- define a series of putatively causative polymorphisms for each QTL (King, *et al*. 2012a). At the present we phase only QTL identified in the inbred, OreRx cross, and Iso1x cross. Founder alleles

are not phased in the RIX cross due to the presence of population structure within the RILs, and its likelihood to cause mis-estimation of founder means at QTL in the RIX cross.

Results:

Variation in starvation resistance among crosses to DSPR founder lines:

Phenotypes of the F_1 progeny are more resistant than the inbred DSPR founder lines (t-test: Founders and OreR x Founder Cross: p=0.041, Founders and Iso1 x Founders Cross: p=0.002). F_1 progeny from the vast majority of the crosses were found to be more starvation resistant than their midparent value (expected under additivity), revealing the existence of substantial heterosis for starvation resistance (fig.2). Furthermore, the phenotypes of DSPR founders and that of the F_1 progeny of those founders are not significantly correlated (Founders and OreR x Founder Cross r^2 =0.099, p=0.7247; Founders and Iso1 x Founder Cross: r^2 =0.300, p=0.2765), also suggesting a deviation from additivity for *D. melanogaster* starvation resistance. Interestingly, the correlation between the phenotypes of the two crosses (Founders x OreR and Founders x Iso1) is strong (r^2 =0.620, p=0.014), and suggests that outcrossing provides a general positive effect on starvation resistance.

Starvation resistance among crossing designs utilizing the DSPR:

Starvation resistance phenotypes of F_1 progeny of RILs and isogenic lines are significantly correlated with the phenotype of the RIL used in the cross (Inbred/Iso1x: r^2 =0.354, p < 2.2e-16; Inbred/OreRx: r^2 =0.380, p < 2.2e-16). This highly significant

correlation was not observed in the crosses of the founder lines and isogenic strains, likely due to the vastly smaller number of founder lines as compared to pA RILs. The F_1 phenotypes of the RIL and isogenic line crosses are significantly correlated with each other (Iso1x/OreRx: r^2 =0.321, p < 2.2e-16). Together, these results indicate that there is a positive relationship between a inbred RIL phenotype and the phenotype of its outcrossed progeny.

The mean of the phenotypic distribution of the OreRx cross is substantially higher than any other experimental design (all pairwise t-tests between the OreRx and the 3 other designs p < 2.2e-16; fig. 3), potentially suggesting the *OreR* genotype possesses allelic variants that confer some resistance to starvation. Interestingly, the inbred RIL phenotypic distribution possesses both the most strongly and weakly starvation resistant animals, pointing toward a potential moderating effect of heterozygosity on RILs with extreme phenotypes.

QTL for starvation resistance among crossing designs:

Among the four mapping experiments utilizing females derived from pA RILs we discovered 21 QTL surpassing a 95% confidence threshold (5 of which are reported in Chapter 2) (see triangles, fig. 4). Seven of the QTL are shared between two crosses, with one QTL being identified among all four mapping experiments (colored triangles, fig. 4). The remaining 13 QTL were identified in only one mapping experiment (table 1).

Notably, 4/5 QTL originally identified in the pA RILs were verified through re-discovery in at least one heterozygous mapping cross. The single QTL that was not remapped (QTL pA3.1, table 1 & fig.4 magenta) approaches the genomewide

significance threshold in the OreRx mapping experiment (peak LOD = 6.89, threshold = 7.0). Having an *a priori* expectation that a QTL exists at that location from the mapping results obtained in the inbred RIL mapping experiment, one may consider this QTL to be verified. Correspondingly, we name the suggestive peak in the OreR cross ORAf_3.S ('S' for suggestive). Indeed, previous work has shown that the effect of a QTL is often over-estimated when a QTL is initially identified, the so-called 'winner's curse' (Beavis, 1998), making a convincing argument that a suggestive peak identified in a second mapping experiment is strong evidence to support the original QTL finding.

The QTL identified among inbred pA RILs, defined as a 2-LOD reduction from the peak intensity, encompass an average of 970kbp (1.98cM), and include a mean of 132.8 protein-coding genes (range: 36-202) within a 2-LOD reduction from the QTL peak (table 1). Crosses utilizing pA RILs resolve QTL slightly better, including an average 2-LOD distance of ~753kbp (1.83cM), capturing ~94 genes on average (table 1). Individual QTL identified in crosses contain as few as 25 protein-coding genes, and as many as 396 for a QTL located at the centromere of chromosome 3.

Phasing of QTL identified in backcross designs and inbreds:

QTL phasing utilizes the phenotypic effects of founder alleles at each QTL to assess which founders share the same causative allele(s). Founders contributing similar effects at a QTL are identified as sharing a phase (e.g. having a 'resistant' allele, as opposed to a 'susceptible' allele). Here I discuss the phasing results for QTL that were discovered in multiple mapping experiments.

Four QTL were identified in multiple QTL mapping crosses that we are able to phase (fig. 5a,b,c,d). QTL for each cross are phased independently at the physical position corresponding to the localized peak LOD score. For the single QTL that was identified in all mapping experiments (QTL pA2.4/q_ORAf_2.4/q_SSAf_2.3) we identify a surprising pattern of QTL phases (fig. 5a). In the three crosses, founder line AB8 is identified as harboring the 'resistant' allele, while founder A4 is among the contributors to the 'susceptible' founder allele class. Defying a simplistic explanation, founder lines A5 and A6 are both identified as contributing the 'susceptible' allele in the inbred RIL experiment (the allele of A5 is also 'susceptible' in the OreRx cross), while both are considered to deliver the 'resistant' allele in the Iso1x cross. It is unlikely that the founder effects were poorly estimated in any mapping experiment, as a large number of RILs were used to phase the founder lines across all three mapping experiments (range = 45-305 RILs). The plethora of RILs used to phase each founder genotype should yield a high-confidence estimate of each founder allele's true effect. This result indicates that the phenotypic contribution of the alleles conferred by founder A5 and A6 at this QTL is likely dependent upon the genetic background in which they are assayed.

Two of the QTL that were identified in two phasable crosses have similar phasing patters across experiments. As seen in figure 5b, founder lines at the QTL pA2.2/SSAf_2.2 (cyan triangles fig.4) are phased identically in both the inbred RILs and Iso1x cross. The purple colored QTL identified in fig. 5c (ORAf_3.2/SSAf_3.1) is also phased similarly in each cross, though each mapping design fails to phase one founder. This pattern of allelic variation appears to be consistent with a QTL that has consistent effects in multiple mapping crosses. The third shared QTL, pA3.1/ORAf_3.S, (magenta

triangles fig. 4) is more complex (fig. 4d). In both the inbred RILs and in the OreRx cross, founder A7 clearly possesses an allele that is much more resistant than are the alleles of any other phased founders. Among founders A2-A6, however, there are two allelic phases, with the weakest allele being attributed to founder A5 in the inbred RIL experiment, but to founder A3 in the OreRx cross (fig. 5d black bars).

Discussion

In this study we begin to critically examine the genetic architecture of an important life-history trait -- starvation resistance -- in multiple genetic backgrounds using an otherwise similar methodological strategy. We assayed starvation resistance in adult female *D. melanogaster* as both a series of homozygous inbred genotypes, and in three sets of heterozygous outbred collections of genotypes. Similar to other quantitative genetic studies of life history traits in *D. melanogaster* (Mackay, et al. 2012, Harbison, et al. 2004, Vieira, et al. 2000), we find starvation resistance to be a highly polygenic trait, with multiple loci of small to moderate effect influencing this phenotype. Comparing the mapping results obtained for starvation resistance phenotypes in populations of inbred lines and crosses between lines, our results reveal an unexpectedly complex and genetic architecture, while suggesting that the genetic contribution to starvation resistance is often background specific.

We hypothesized that several, if not many QTL identified among inbred lines were generated due to a generalized reduction in organismal fitness caused by inbreeding. Directly testing this hypothesis by mapping QTL in heterozygous animals, we discovered that all five QTL identified among inbred females were validated through

re-discovery in an alternative mapping design. Indeed, the three QTL mapped in the inbred RIL experiment (inbreeding coefficient '*f*' =1) and in crosses between RILs and the isogenic strains (*f* = 0) cannot be due to inbreeding depression. The two remaining female pA RIL QTL were re-discovered in the RIX crossing scheme are also unlikely to be caused by inbreeding depression (RIX crosses *f* = 0.125). Knowledge that the QTL identified in the inbred RIL experiment are specific to starvation resistance phenotypes, and not merely observed due to inbreeding depression increases our confidence in pursuing a mechanistic understanding of causative genes under these QTL. Given the fraction of QTL identified by direct assay of homozygous RILs that are not likely to be inbreeding depression loci, we can be confident that using panels of inbred lines to uncover the genetic basis of complex is a valid experimental practice.

Despite the success of replicating all of the QTL identified in inbred pA RILs, a perplexing observation remains: Thirteen of the 21 QTL were identified in only one mapping cross. There are several possibilities that could account for such an observation. We discuss each possibility below, though it should be noted that with the present data we cannot determine the precise cause(s) of the large number of unique QTL.

First, the QTL discovered among inbred pA RILs contribute relatively small effects to starvation resistance -- each explaining between 4.4% and 6.7% of the estimated heritability of starvation resistance in female pA RILs. These effect sizes remain fairly constant among cross designs, as none of the 13 cross-specific QTL, have a phenotypic effect greater than 5.5%. This poses a technical challenge, as the power to detect a QTL is directly related to its effect. For example, simulations using the full set

of pA inbred RILs reveal that we have 84% power to detect a 5% QTL (King, et al. 2012b). Detecting small effect QTL in crosses is more difficult; power drops to 68% to identify a 5% QTL among crosses between pA and pB. Similarly, our RIX cross design (pA x pA) has less power than using inbred pA RILs directly. The two backcross design experiments also suffer from reduced statistical power, especially in the case of Iso1x cross, where n=779 (rather than the full panel of 861 RILs), and only one replicate vial was assayed for starvation resistance (as compared to two vials for all other experiments). These results indicate that it is likely we stochastically fail to identify true QTL due to the statistical challenges of mapping QTL of small effect.

Second, QTL identified in only a single backcross experiment may be masked by a dominant allele in either of the isogenic strains serving as the recurrent parent. For example, nine QTL were identified in the OreRx cross. Of those, only two were re-identified in the Iso1x cross, despite an identical crossing design. This result is consistent with the presence of dominantly-acting alleles at causative genes in the *Iso1* genome, since such alleles will mask the differences among DSPR founder alleles at the QTL. This result is distinct from that expected if any of the alleles segregating within the DSPR exhibit substantial dominance. Any dominant QTL identified in the pA inbred lines should be replicated in the RIX design, but it may not be replicated in a backcross experiment, where all test individuals share an identical complement of maternally-derived chromosomes. Ultimately, only by cloning the causative site(s) and/or assaying each QTL allele in many genetic backgrounds will we be able to identify the level of dominance/recessiveness of each QTL allele.

Third, the effect of 'cryptic' genetic variation can cause the appearance or disappearance of QTL in a novel genetic background (reviewed in Gibson and Dworkin, 2004). Crossing the DSPR RILs to any other strain introduces QTL alleles into a novel genetic background. This cryptic genetic variation can be discovered through its altering of the effects of both 'resistant' and 'susceptible' QTL alleles, increased genetic variance between the two QTL alleles (decanalization), or epistatic interactions in which the effect of one of the two QTL alleles is altered by the new genetic background.

In the present experiment we have some evidence for the effects of cryptic genetic variation acting on starvation resistance QTL. For two QTL initially discovered among inbred pA RILs (pA2.4 and pA3.1) phasing results indicate that allelic effects differ across genetic backgrounds. For QTL pA2.4 founders A5 and A6 share a 'susceptible' allele, while in QTL SSAf_2.3 they are both phased as 'resistant' (fig. 4a). Not only does this alteration in allelic effects imply the presence of cryptic genetic variation acting on this allele, but the disparate phasing also implies that these two founders share an allele not found in any other pA founder lines. A similar phenomenon is observed in QTL pA3.1/ORAf_3.S, in which founder A5 is phased differently in the two mapping designs (fig. 4d). In either case, determining the precise cause of these interesting shifts in allelic effects will require additional study.

An elegant solution for a thorough characterization of the genetic architecture of important life-history traits -- and indeed any QTL or association -- involves investigating the allelic effects of each QTL across many genetic backgrounds, essentially extending our work to additional crossing designs. Nevertheless, such experiments can only give tangential information on the levels of additive, dominance, and epistasis acting on a

phenotype. A useful first step in precisely delineating the contribution of these phenomena would be to carry out a diallel among the DSPR founder lines. This crossing scheme would provide us with valuable estimates of the additivity and dominance existing for starvation resistance among the founding population of the DSPR (Lynch and Walsh 1998). Recently developed analytical platforms (Greenberg, et al. 2010) could additionally be used to perform a sparse diallel (modified from a full all-by-all round-robin crossing design) on the set of DSPR RILs. The large number of recombinant genomes created in this way would provide excellent estimates of the genetic control mechanisms underlying starvation resistance and help to resolve the precise mode of action of causative alleles affecting the phenotype.

Literature Cited

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al. 2000. The genome sequence of *Drosophila melanogaster. Science* **287**: 2185-2195

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker E.U., Cresko, W.A., and Johnson, E.A. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* **3:** e3376.

Beavis, W.D. 1998. QTL analyses: power, precision, and accuracy, pp. 145–162 in *Molecular Dissection of Complex Traits*, edited by A. H. PATERSON. CRC Press, New York.

Broman K.W. and Sen S. 2009. *A Guide to QTL Mapping with R/qtl*. Springer, New York.

Churchill, G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. *Genetics* **138:** 963–971

Dworkin, I. and Gibson, G. 2004. Uncovering cryptic genetic variation. *Nature Reviews Genetics* **5**: 681-690

Forbes, S.N., Valenzuela, R.K., Keim, P., and Service, P.M. 2004. Quantitative trait loci affecting life span in replicated populations of *Drosophila melanogaster*. I. composite interval mapping *Genetics* **168**: 301–311

Greenberg, A.J., Hackett, S.R., Harshman, L.G., and Clark, A.G. 2010. A hierarchical Bayesian model for a novel sparse partial diallel crossing design. *Genetics* **185**: 361–373

Harbison, S.T., Yamamoto, A.H., Fanara, J.J., Norga, K.K., and Mackay, T.F.C. 2004. Quantitative trait loci affecting starvation resistance in *Drosophila melanogaster*

Genetics **166**: 1807–1823

Harshman L.G., and Schmid. J.L., 1998. Evolution of starvation resistance in *Drosophila melanogaster*: Aspects of metabolism and counter-impact selection. *Evolution* **52**(6): 1679-1685

King, E.G., Merkes, C.M., McNeil, C.L., Hoofer, S.R., Sen, S., Broman, K.W., Long, A.D., and Macdonald, S.J.. 2012a. Genetic dissection of a model complex trait using the *Drosophila* Synthetic Population Resource. *Genome Research* (Online before print, April 2012)

King, E.G., Macdonald, S.J., and Long, A.D. 2012b. Properties and power of the *Drosophila*
Synthetic Population Resource for the routine dissection of complex traits. *Genetics* (Epub ahead of print Apr 13, 2012)

Leips, J., Gilligan, P., and Mackay, T.F.C. 2006. Quantitative trait loci with age-specific effects on fecundity in *Drosophila melanogaster*. *Genetics* **172**:1595-605.

Lindsley, D. L. and Zimm, G. 1992. The genome of *Drosophila melanogaster*. New York: Academic Press.

Mackay, T.F., Richards, S., Stone, E.A., Barbadilla, A., Ayroles, J.F., Zhu, D., Casillas,

S., Han, Y., Magwire, M.M., Cridland, J.M., et al. 2012. The *Drosophila melanogaster* genetic reference panel. *Nature* **482:** 173–178

McNeil, C.L., Bain, C.L., and Macdonald, S.J. 2012. Genetic Dissection of Starvation Resistance among Inbred Lines of the *Drosophila* Synthetic Population Resource (in preparation for *Genetics*).

Morgan, T.J., and Mackay, T.F.C. 2006. Quantitative trait loci for thermotolerance phenotypes in *Drosophila melanogaster. Heredity* **96**: 232–242

Nuzhdin, S.V., Khazaeli, A.A., and Curtsinger, J.W. 2005. Survival analysis of life span quantitative trait loci in *Drosophila melanogaster. Genetics* **170**: 719–731

Pasyukova, E.G., Vieira, C., and Mackay, T.F.C. 2000. Deficiency mapping of quantitative trait loci affecting longevity in *Drosophila melanogaster. Genetics* **156**: 1129-1146

Rose, M.R., and Charlesworth, B. 1981a Genetics of life history in *Drosophila melanogaster.* I. Sib analysis of adult females. *Genetics* **97**: 175-186

Rose, M.R., and Charlesworth, B. 1981b. Genetics of life history in *Drosophila melanogaster.* II. Exploratory selection experiments. *Genetics* **97**: 187-196

Tsaih, S.W., Lu, L., Airey, D.C., Williams, R.W., and Churchill, G.A. 2005. Quantitative trait mapping in a diallel cross of recombinant inbred lines. *Mammalian Genome*. **16**(5): 344-55.

Vieira, C., Pasyukova, E.G., Zeng, Z.B., Hackett, J.B., Lyman, R.F., and Mackay, T.F.C. 2000 Genotype-Environment Interaction for Quantitative Trait Loci Affecting Life Span in *Drosophila melanogaster. Genetics* 154:213-227

Wilson, R.H., Morgan, T.J., and Mackay, T.F.C. 2006. High-resolution mapping of quantitative trait loci affecting increased life span in *Drosophila melanogaster. Genetics* **173**: 1455–1463

Figures and Tables

Table 1 (next two pages). Starvation resistance QTL among mapping designs. Each QTL shared among crosses is color coded as in Figures 4 and 5.

[a] QTL are numbered using their population (pA) or cross, chromosome number, and order of peak position (e.g. ORAf_2.1 is from the OreR x RIL cross, chromosome 2, first QTL)

[b] The physical position (bp) of each QTL peak for flybase release FB2012_04 (July 6th, 2012)

[c] The size of each QTL 2-LOD support interval in Mb

[d] The number of founder means that are phased as the minor allele over the total number of founder means phased at a given QTL. At present we do not phase QTL in the RIX crossing design, and therefore cannot assign the frequency of QTL alleles in the RIX cross.

[e] The number of protein-coding genes within each 2-LOD interval.

Table 1

Mapping Design	QTL Name[a]	Color	Chr	Ppos [b]	Size (Mb)[c]	LOD Peak	Minor Allele Freq.[d]	Gene Number[e]
RIL	pA2.2		2R	5240000	1.37	9.57	1/5	202
RIL	pA2.3		2R	13380000	0.90	7.57	1/3	146
RIL	pA2.4		2R	14710000	0.58	7.22	1/5	94
RIL	pA3.1		3L	3650000	0.52	10.49	1/3	36
RIL	pA3.2		3R	1180000	1.48	11.12	2/4	186
RIX	RRAf_1.1		X	5810000	0.27	9.20	NA	36
RIX	RRAf_1.2		X	8110000	0.50	8.26	NA	69
RIX	RRAf_1.3		X	11960000	0.40	7.85	NA	25
RIX	RRAf_1.4		X	17960000	1.62	7.45	NA	166
RIX	RRAf_2.1		2L	3680000	0.43	7.36	NA	59
RIX	RRAf_2.2		2L	12020000	0.36	8.28	NA	64
RIX	RRAf_2.3		2L	15430000	1.12	7.10	NA	94
RIX	RRAf_2.4		2R	13240000	0.75	8.51	NA	112
RIX	RRAf_2.5		2R	14500000	1.19	9.00	NA	166
RIX	RRAf_3.1		3L	1680000	0.50	7.75	NA	89
RIX	RRAf_3.2		3L	5900000	1.02	7.63	NA	127
RIX	RRAf_3.3		3L	24360000	2.25	16.87	NA	396

Table 1 (continued)

Mapping Design	QTL Name[a]	Color	Chr	Ppos [b]	Size (Mb)[c]	LOD Peak	Minor Allele Freq.[d]	Gene Number[e]
OreR	ORAf_2.1		2L	5520000	0.55	7.07	3/7	53
OreR	ORAf_2.2		2L	19260000	1.24	8.16	2/4	166
OreR	ORAf_2.3		2R	4650000	0.64	8.51	1/5	97
OreR	ORAf_2.4		2R	14430000	0.58	7.41	2/5	98
OreR	ORAf_3.S		3L	3420000	0.36	6.88	1/5	36
OreR	ORAf_3.1		3L	5930000	1.71	8.29	2/5	175
OreR	ORAf_3.2		3L	10270000	0.48	7.38	1/3	24
OreR	ORAf_3.3		3R	22700000	0.66	8.44	2/5	71
OreR	ORAf_3.4		3R	24480000	0.74	7.34	2/6	45
Iso1	SSAf_1.1		X	14090000	0.40	6.82	1/3	44
Iso1	SSAf_2.1		2L	15580000	0.46	8.85	3/6	29
Iso1	SSAf_2.2		2R	5950000	0.33	7.63	1/5	64
Iso1	SSAf_2.3		2R	14510000	0.62	6.82	1/5	100
Iso1	SSAf_3.1		3L	10150000	0.39	7.75	1/3	45

Figure 1

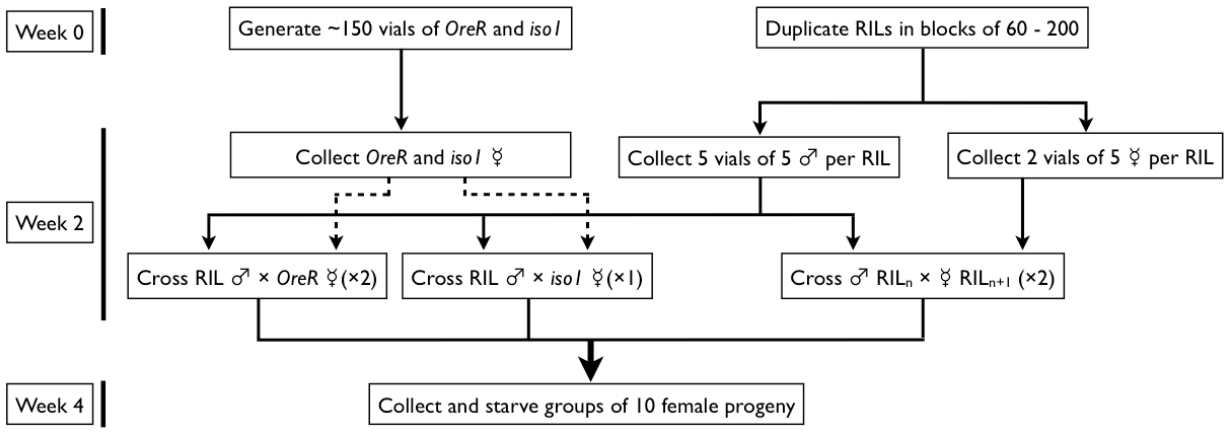


Fig. 1: Schematic of crosses performed in this experiment.

Crosses were performed among the ~850 pA RILs and between each RIL and two isogenic reference strains in experimental blocks of 60-200 RILs. In all cases vials of 5 - 10 virgin females were paired with vials containing 5 males. F1 progeny of the RIL x RIL and RIL x isogenic strain crosses were assayed for starvation resistance.

Figure 2

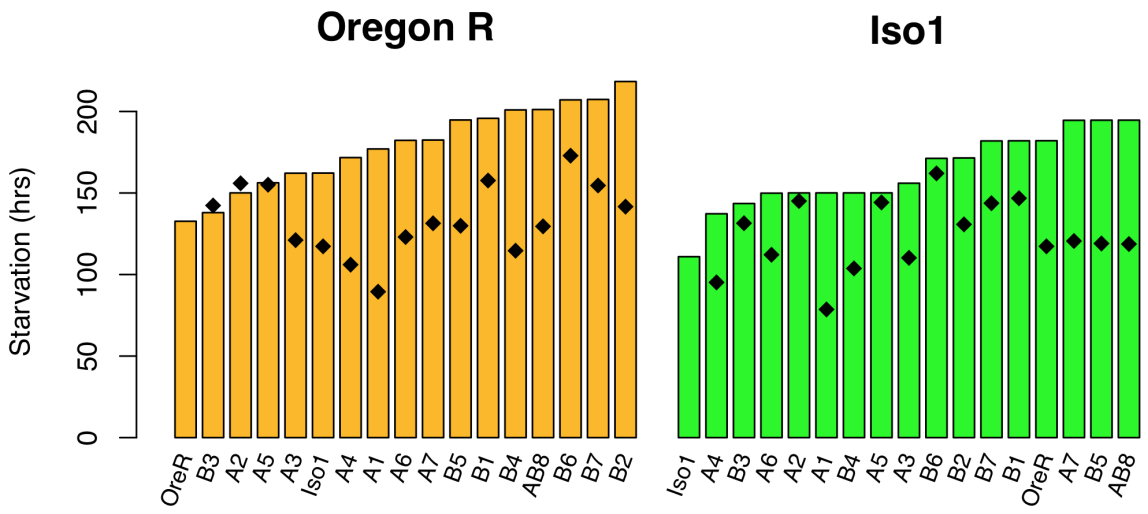


Fig. 2: Phenotypic distribution of the progeny of DSPR founders and isogenic reference strains used in this study.

The median starvation-induced death time are shown for crosses between the DSPR founder lines and two isogenic reference strains. Black diamonds for each cross indicate the expected midparent value under additivity for starvation resistance phenotypes. Progeny of crosses to both isogenic reference strains show substantial heterosis for starvation resistance, as F1 progeny typically outperform predicted F1 phenotypes.

Figure 3

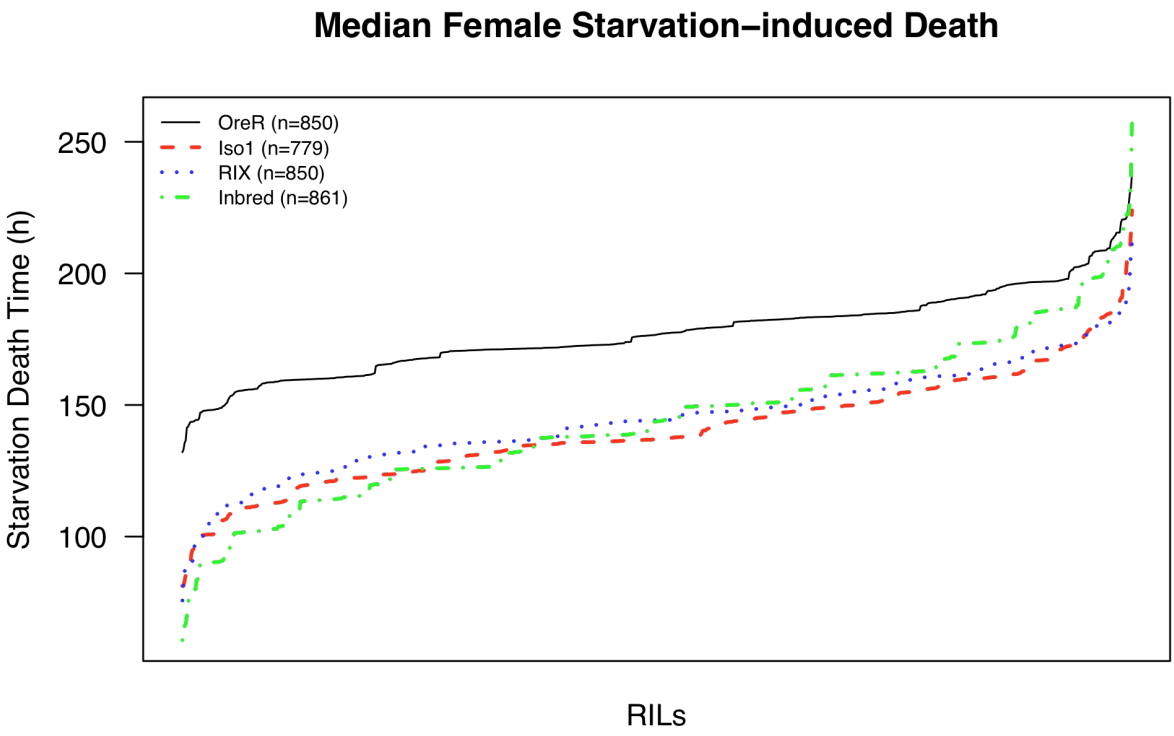


Fig. 3: Phenotypic distributions of the progeny of DSPR founders and isogenic reference strains used in this study.

The median starvation-induced death times for each cross and the DSPR inbred lines are ordered from least to most resistant. Phenotypes of crosses utilizing the RIX design (blue) and RIL x Iso1 backcross (red) are distributed much like the inbred RILs (green). Progeny of the RIL x OreR crosses (black) are substantially more resistant to starvation stress than animals from other experimental designs.

Fig. 4: Starvation Resistance QTL among the experimental designs. (next page) QTL for starvation resistance are highlighted by triangles on the x-axis. White triangles indicate QTL that are indentified in a single mapping experiment. The single QTL that crosses our 95% genomewide significance threshold in all mapping crosses is highlighted in red. QTL represented by triangles of other colors (blue, cyan, green, orange, purple, and yellow) represent QTL that were discovered in two mapping designs.

Figure 4

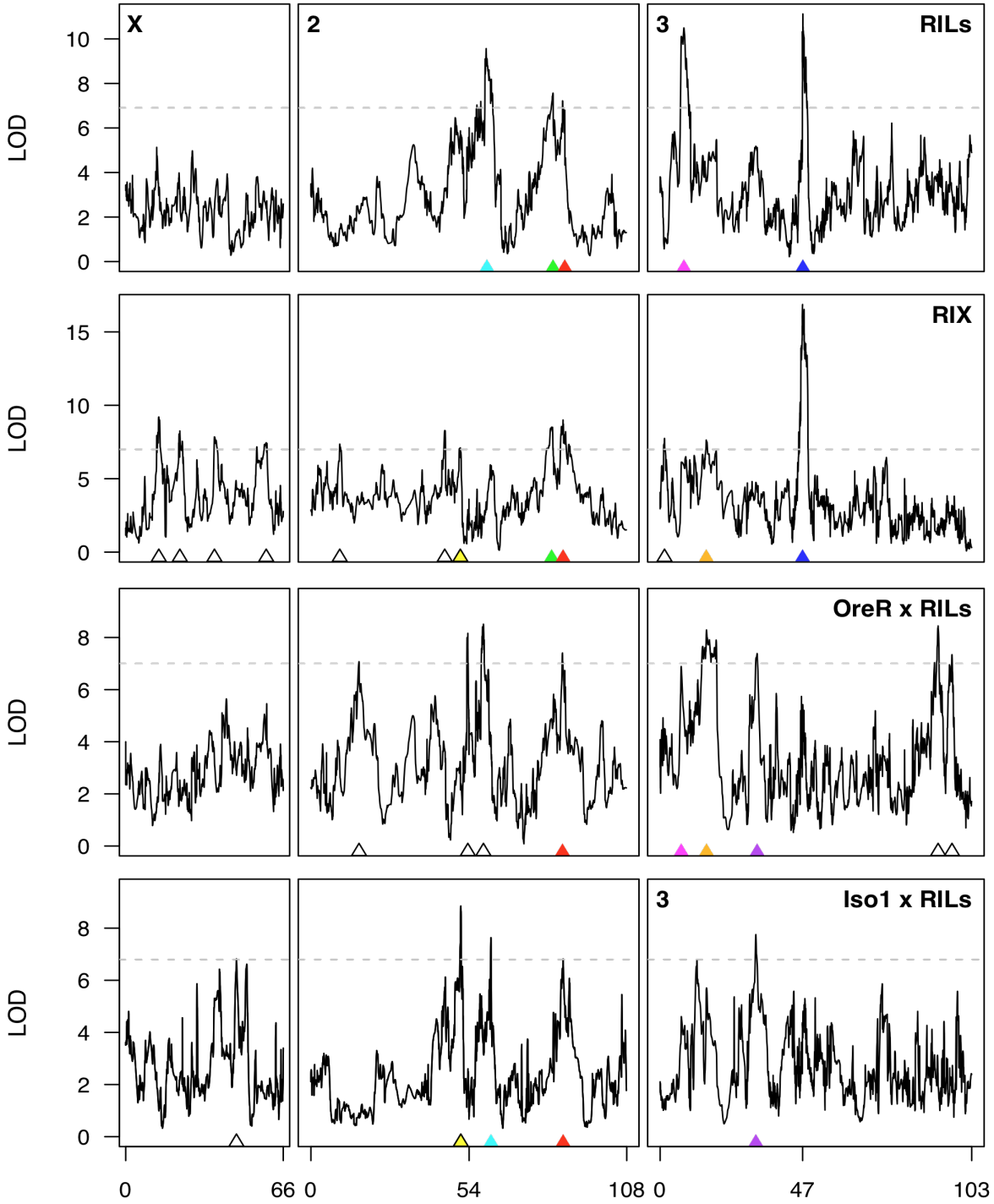


Fig. 5: Founder contributions to the QTL shared among all crosses.

Mean starvation resistance for founder haplotypes at QTL identified in the inbred RILs and backcross designs. Bars are color coded to match QTL in table 1 and fig. 4. In each case, the 'resistant' allele of each QTL is colored as in fig. 4, with bars representing the 'susceptible' allele being black (white or grey bars represent unphased founders). In the case of Fig. 4d, light-pink bars indicate founder alleles of intermediate phase.

Figure 5a

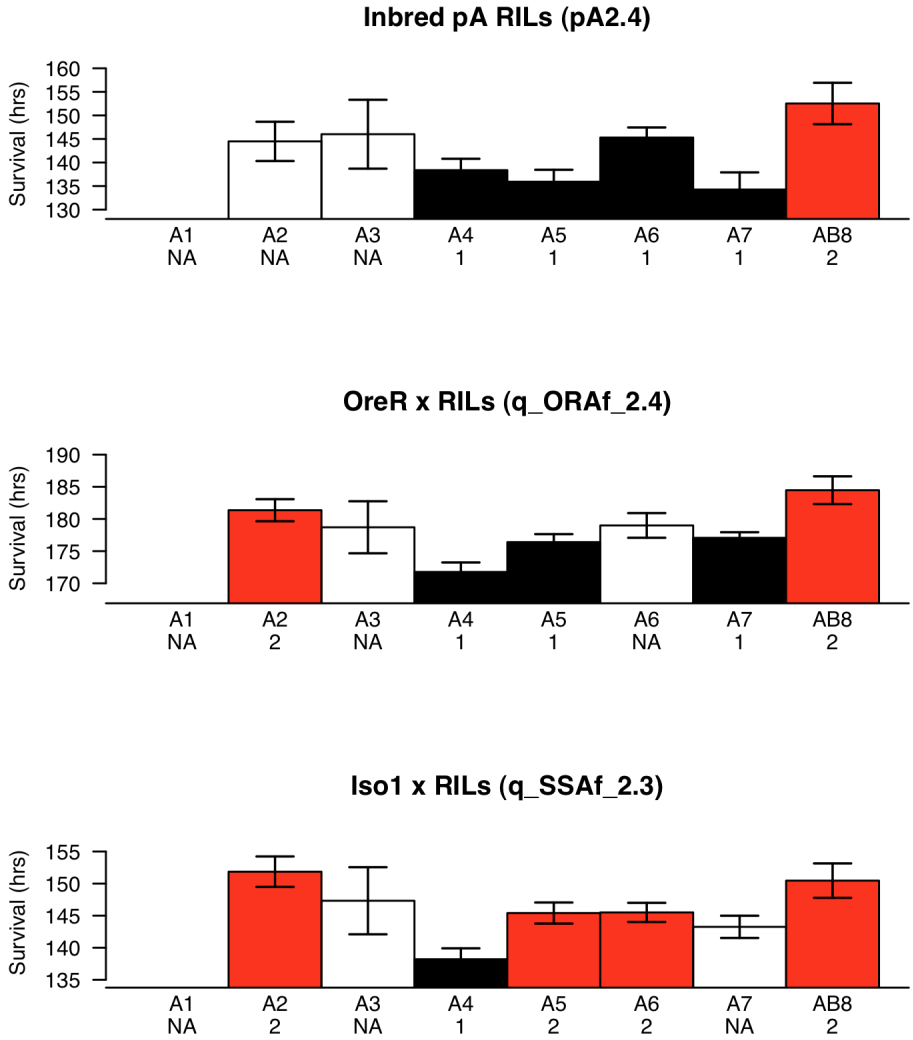


Figure 5b

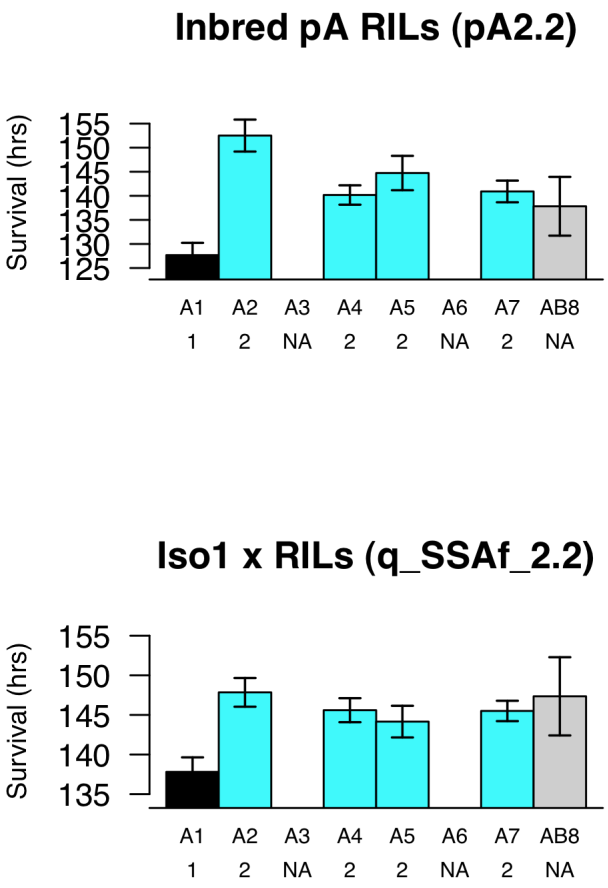

Inbred pA RILs (pA2.2)
Survival (hrs)
155
150
145
140
135
130
125
A1 A2 A3 A4 A5 A6 A7 AB8
1 2 NA 2 2 NA 2 NA
Iso1 x RILs (q_SSAf_2.2)
Survival (hrs)
155
150
145
140
135
A1 A2 A3 A4 A5 A6 A7 AB8
1 2 NA 2 2 NA 2 NA

Figure 5c

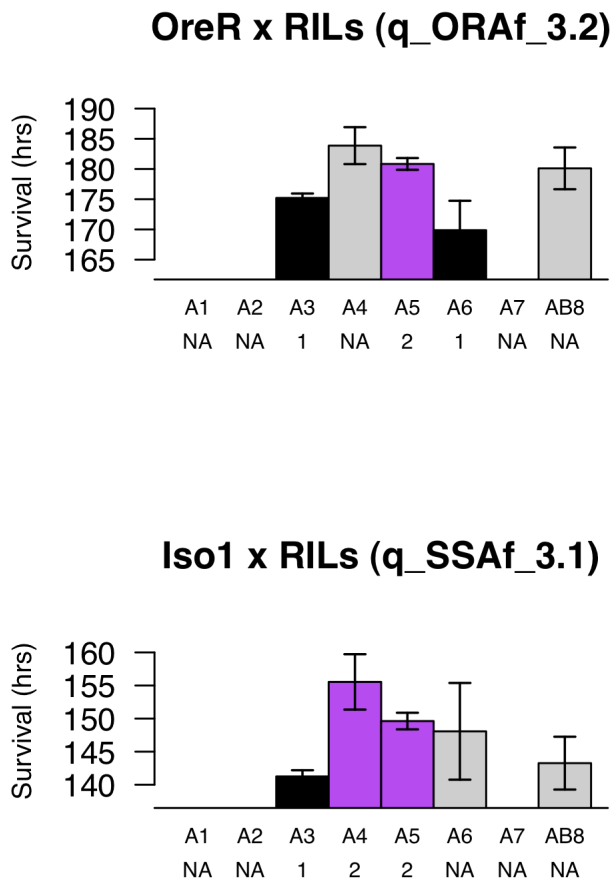
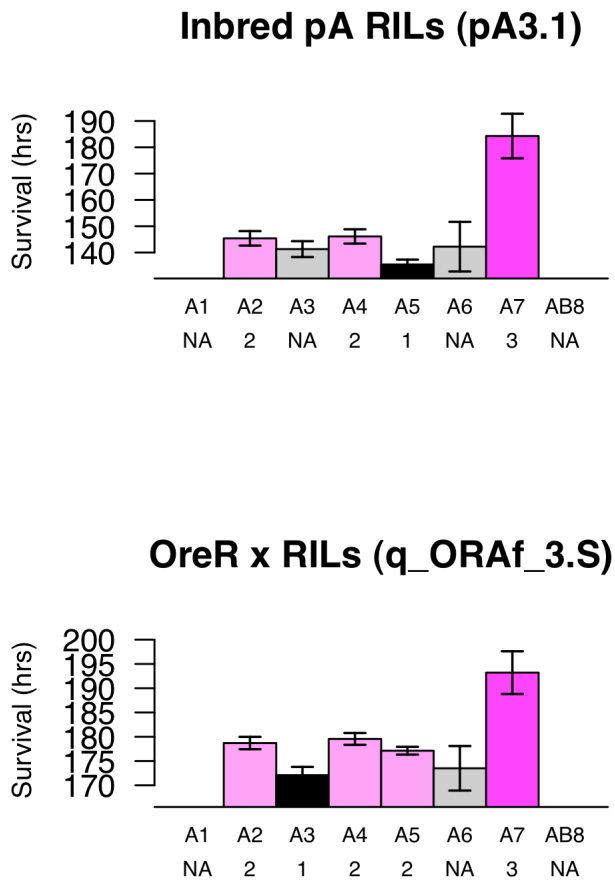

OreR x RILs (q_ORAf_3.2)
Survival (hrs)
190
185
180
175
170
165
A1 A2 A3 A4 A5 A6 A7 AB8
NA NA 1 NA 2 1 NA NA
Iso1 x RILs (q_SSAf_3.1)
Survival (hrs)
160
155
150
145
140
A1 A2 A3 A4 A5 A6 A7 AB8
NA NA 1 2 2 NA NA NA

Figure 5d


Inbred pA RILs (pA3.1)
Survival (hrs)
190
180
170
160
150
140
A1 A2 A3 A4 A5 A6 A7 AB8
NA 2 NA 2 1 NA 3 NA
OreR x RILs (q_ORAf_3.S)
Survival (hrs)
200
195
190
185
180
175
170
A1 A2 A3 A4 A5 A6 A7 AB8
NA 2 1 2 2 NA 3 NA

Conclusion

Historically, the understanding of genetic contributions quantitative life history traits has been hampered by their polygenic nature as well as the rather large contribution of environmental variation acting on these phenotypes. Though correlations between various physiological life history traits have been heavily studied, there has been relatively little attention paid to identifying the effects and frequencies of genetic variants affecting life history traits in natural populations of animals. Since there is relatively little known about the quantitative genetics underlying these traits, we sought to utilize the elite model system *Drosophila melanogaster* to better understand these important traits.

This study focuses on two traits: the morphology of the posterior lobe of the genital arch in male *D. melanogaster,* and starvation stress resistance. Both traits are likely subject to direct natural selection. The shape of the posterior lobe is hypothesized to influence male mating success, while the ability to withstand brief periods of nutritional stress is likely may be likely to increase the total number of progeny a single *D. melanogaster* can produce.

The observation that male genitalia diverge more rapidly than other morphological traits during evolution is taxonomically widespread, and likely due to some form of sexual selection. One way to elucidate the evolutionary forces acting on these traits is to detail the genetic architecture of variation both within and between species, a program of research that is considerably more tractable in a model system. *Drosophila melanogaster* and its sibling species *D. simulans*, *D. mauritiana*, and *D. sechellia* are morphologically distinguishable only by the shape of the posterior lobe, a

male-specific elaboration of the genital arch. We extend earlier studies identifying QTL (quantitative trait loci) responsible for lobe divergence across species, and report the first genetic dissection of lobe shape variation within a species. Using an advanced intercross mapping design we identify three autosomal QTL contributing to the difference in lobe shape between a pair of *D. melanogaster* inbred lines. The QTL each contribute 4.6-10.7% to shape variation, and two show a significant epistatic interaction. Interestingly, these intraspecific QTL map to the same locations as interspecific lobe QTL, implying some shared genetic control of the trait within and between species. As a first step towards a mechanistic understanding of natural lobe shape variation, we find an association between our QTL data and a set of genes that show sex-biased expression in the developing genital imaginal disc (the precursor of the adult genitalia). These genes are good candidates to harbor naturally-segregating polymorphisms contributing to posterior lobe shape.

Populations of animals commonly encounter intermittent periods of environmental stress in nature, making quantitative differences in an individual's ability to resist environmental challenges a target for natural selection. One major stress that animals may encounter in nature is the lack of nutritious food. While numerous studies have attempted to characterize the genetic contribution to starvation-resistance phenotypes in *Drosophila,* there has been relatively little success in identifying and validating genes that lead to quantitative variation in starvation resistance. To more thoroughly investigate the nature of the genetic variation controlling starvation resistance, we map QTL for starvation resistance using the *Drosophila* Synthetic Population Resource (King, *et al.* 2012). The DSPR allows for simultaneous estimates

of QTL effect and frequency, while the large number of recombinant inbred lines (RIL) provides much needed statistical power to repeatedly detect QTL of small effects. We starved 75,545 flies from 1725 RILs of the DSPR, facilitated by implementing high-throughput phenotyping strategies utilizing barcoded vials and semi-automated data collection. In the two populations of eight-way DSPR recombinant inbred lines (RILs) we identify at least 17 autosomal QTL. All QTL are specific to a single population of the DSPR and five are shared between males and females. The 2-LOD confidence interval around each QTL mapped to very small genetic and physical intervals (mean = 1.28cM and 0.91Mb) and encompasses ~115 genes per QTL peak.

To explore physiological correlates for our starvation resistance phenotype, we carried out several additional experiments using lines from the tails of the starvation resistance distribution. Reasoning that flies that displayed increased activity while under starvation conditions would more quickly deplete their energy reserves, we assayed a subset of the most and least starvation resistant RILs in the *Drosophila* Activity Monitor (DAM). DAM activity data (generated under starvation conditions) for animals corresponding to the tails of the starvation resistance distribution showed a trend of increased activity in RILs that are more susceptible to starvation. This observation suggests that behavioral traits may play a role in starvation resistance phenotypes, however, it does not exclude other explanations for quantitative variation in starvation resistance. To examine the possibility that starvation resistance might be associated with a general stress response, we investigated whether the ability to withstand starvation resistance is correlated with resistance to desiccation stress. Assaying both exceptionally starvation resistant and susceptible RILs we demonstrate that lines of

strong starvation resistance phenotypes are more desiccation resistant than those that succumb quickly to starvation. This result indicates that starvation resistance is partially controlled by a generalized stress response mechanism and that other stress-related life history traits could be under similar genetic control.

Finally, we attempted to identify potentially causative alleles by combining data derived from our powerful QTL mapping experiment with the high-resolution association mapping results of a recent GWAS for starvation resistance (Mackay, *et al.* 2012). Surprisingly, only 9/203 SNPs implicated by GWAS (at an arbitrary threshold $p < 10^{-5}$) reside within a QTL. Conditioning on the fact that we mapped 17 QTL and thus had *a priori* evidence for the existence of causative loci within these intervals, we performed a targeted association study: By reducing the number of tests the threshold for statistical significance is also reduced. Though this approach greatly reduced the number of statistical tests performed, the analysis identified only a single association after Bonferroni correction.

The lack of coincidence between our QTL mapping and the GWAS data could be explained in several ways. It is possible that the genetic contribution to starvation resistance is highly population-dependent. Since the GWAS panel was derived from wild-caught flies isolated in a single location and the DSPR was founded by isolates of worldwide origin, it is possible that we are not sampling similar sets of functional genetic variants. More likely however, is that GWAS have low power to detect either rare variants (Pritchard, *et al*. 2001) or common alleles of small effect (Long and Langley, 1999). As our QTL mapping demonstrates that the vast majority of QTL contribute relatively little to the starvation resistance phenotype, it is likely that GWAS, at least as

implemented in the ~180 lines of the DGRP (Mackay, *et al.* 2012), is not an effective strategy for the identification of alleles influencing life history traits in *Drosophila.*

Although the DSPR provides a number of advantages over prior linkage mapping and common association mapping approaches, there are two major concerns with relying on genetic mapping within inbred lines. First, inbred, homozygous genotypes are clearly not representative of naturally-derived flies. Second, mapping directly in inbred lines allows for the identification of QTL in only a small subset of genetic backgrounds, rather than the innumerable permutations of genetic backgrounds obtainable in nature.

In light of these challenges, we mapped QTL for starvation resistance utilizing the DSPR in three outbred genetic environments. Mapping QTL in a series of round robin crosses between lines of the DSPR and backcrosses between DSPR lines and two isogenic strains, we discover a highly complex genetic architecture for starvation resistance, identify many cross-specific QTL, all of which have small effects. Notably, each QTL originally mapped in inbred lines of the DSPR was replicated in at least one cross, indicating that QTL mapped among inbred lines are specific to starvation resistance and not due to inbreeding depression.

Drastically variable effects of genetic background on organismal phenotypes have been known to exist for many years (*e.g*., Threadgill, *et al*. 1995). These background effects hold true for variability in both the penetrance (Dworkin, *et al*. 2009) and expressivity of Mendelian traits. Among quantitative phenotypes, ignoring the potential effects of genetic background can lead to severely misleading consequences (*e.g*., the *Sir2* controversy in *C. elegans,* Burnett, *et al*. 2011, Tissenbaum and Guarente, 2001, Rogina and Helfand, 2004, and Viswanathan, *et al*. 2005). Despite

widespread understanding of the potentially profound effects of genetic background, relatively few empirical studies have attempted to address the consequences of such effects for mapping and understanding the genetic basis of complex trait variation. This is especially true in the study of fitness and life history traits among animals, as their highly polygenic nature and lack of naturally-occurring alleles of strong effect present major challenges in the identification of background-specific effects. While there have been many studies aimed at uncovering the nature of genetic variation of life history traits in general, and starvation resistance in particular, no study had attempted to explore the relative importance of genetic background in these phenotypes.

We devised a genetic mapping scheme that was capable of identifying novel, background-specific QTL, while simultaneously allowing for validation and refinement of QTL for starvation resistance identified within inbred lines. We used this study to help address two major questions. First, does the genetic architecture for starvation resistance vary among genetic backgrounds? Second, are QTL for starvation-resistance identified among inbred lines the product of inbreeding depression? Our strategy for answering these questions centered on phenotyping F_1 progeny of round-robin crosses between RILs and F_1 progeny of crosses between the RILs and two independent, isogenic strains.

Correlating the phenotypes of pA inbred RILs with the F_1 progeny of the OreRx, and Iso1x crosses (each testing RIL genotypes in a different genetic background), we find a weak but significant positive relationship among the phenotypes, (r^2 is between 0.32 - 0.38 for all correlations between Inbreds, OreRx, and Iso1x phenotypes). This weak correlation suggests that while there is some shared genetic control of starvation

resistance among the genetic backgrounds, the effect of genetic background is substantial. In agreement with the phenotypic correlation, comparing QTL identified among RIL, RIX, OreR, and Iso1 mapping designs, only 8/21 QTL were identified in more than one mapping design. While there are several possibilities as to why any individual QTL may be unique to a genetic background, the balance of the evidence points toward a highly background-specific genetic architecture for starvation resistance.

Among the RIL, OreR, and Iso1 mapping crosses we were able to assign QTL 'phases' for the founder alleles at each shared QTL. This phasing allowed us to group founder lines as sharing an allele that is either 'resistant' or 'susceptible' to starvation stress, or of 'intermediate' effect. Of the four QTL that we were able to phase in multiple crosses, we identify two QTL in which founder alleles are phased consistently between crosses, and two QTL in which at least one founder is phased differently. While these phasing results should be interpreted cautiously, they do suggest that while some QTL appear to consistently act in an additive manner, allelic effects at other starvation QTL may be equally likely to be background specific (*i.e*., a fraction of alleles at a locus can have different effects when tested in different backgrounds).

The matter of validating QTL originally identified in inbred lines is by comparison, an easily interpretable result. While only one QTL was identified in all four genetic backgrounds, 4/5 QTL identified among the inbred lines were re-discovered in at least one crossing experiment. The lone QTL that was not re-discovered very nearly surpassed the genomewide significance threshold in the OreR backcross experiment. Conditioning on the *a priori* expectation that a true QTL is localized to this position in the

inbred RIL experiment, we consider this result evidence that we validate this QTL. As all three forms of crossing greatly reduce the inbreeding coefficient (Inbred $f = 1$, RIX $f \sim 0.125$, and OreRx & Iso1x $f = 0$), we can conclude that each QTL identified in the RIL experiment is starvation specific and is unlikely to be a consequence of inbreeding depression.

While a more complete characterization of any QTL for starvation resistance must be accomplished through repeated genetic mapping experiments across multiple genetic backgrounds and ultimately an introgression of the implicated region into other backgrounds, our initial experiment measuring life-history traits in various outcrossed genetic backgrounds is an important first step towards the thorough characterization of this fitness-related trait in *Drosophila.* Furthermore, having recapitulated each of the QTL identified among inbred lines gives us confidence that, as we pursue a molecular understanding of various quantitative life history traits, we are typically targeting our trait of interest, and not false signals due to inbreeding depression. In total, our studies with the DSPR show that starvation resistance is governed by a highly complex genetic architecture composed of many background-specific QTL of small effect. While this complex genetic architecture may preclude the identification of any single causative polymorphism, it is likely to be indicative of the quantitative genetic nature of other life-history traits and highlights the importance of studying complex genetic variation in a variety of genetic backgrounds.

Literature Cited

Burnett, C., Valentini, S., Cabreiro, F., Goss, M., *et al.,* 2011. Absence of effects of *Sir2* overexpression on lifespan in *C. elegans* and *Drosophila. Nature* **477**: 482-486.

Chatterjee, S. S., L. D. Uppendahl, M. A. Chowdhury, P-L. Ip, and M. L. Siegal, 2011. The female-specific Doublesex isoform regulates pleiotropic transcription factors to pattern genital development in *Drosophila*. *Development* **138:** 1099–1109.

King, E.G., Merkes, C.M., McNeil, C.L., Hoofer, S.R., Sen, S., Broman, K.W., Long, A.D., and Macdonald, S.J.. 2012a. Genetic dissection of a model complex trait using the *Drosophila* Synthetic Population Resource. *Genome Research* (Online before print, April 2012).

Long, A.D., and Langley, C.H., 1999. The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Research* **9**:720-731.

Mackay, T.F., Richards, S., Stone, E.A., Barbadilla, A., Ayroles, J.F., Zhu, D., Casillas, S., Han, Y., Magwire, M.M., Cridland, J.M., et al. 2012. The *Drosophila melanogaster* genetic reference panel. *Nature* **482:** 173-178.

Pritchard, J. K., 2001. Are rare variants responsible for susceptibility to complex diseases? *Am. J. Hum. Genet.* **69**: 124–137.

Tissenbaum, H. A. & Guarente, L. 2001. Increased dosage of a *sir-2* gene extends lifespan in *Caenorhabditis elegans*. *Nature* **410**: 227-230.

Viswanathan, M., Kim, S. K., Berdichevsky, A. & Guarente, L. 2005. A role for SIR-2.1 regulation of ER stress response genes in determining *C. elegans* life span. *Dev. Cell* **9**: 605–615.

Rogina, B. & Helfand, S. 2004. *Sir2* mediates longevity in the fly through a pathway related to calorie restriction. *Proc. Natl Acad. Sci. USA* **101**: 15998–16003.